

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.⁷

H04N 5/00

G11B 20/10

[12] 发明专利申请公开说明书

[21] 申请号 01121929.7

[43] 公开日 2002 年 7 月 17 日

[11] 公开号 CN 1359231A

[22] 申请日 2001.6.20 [21] 申请号 01121929.7

[30] 优先权

[32] 2000.12.19 [33] KR [31] 78170/00

[71] 申请人 株式会社考斯默坦

地址 韩国汉城

[72] 发明人 崔元龙 李秉哲 郑尚宪

崔元湜

[74] 专利代理机构 中国专利代理(香港)有限公司

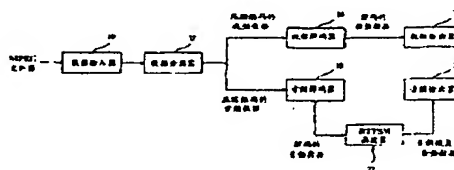
代理人 王 岳 傅 康

权利要求书 5 页 说明书 29 页 附图页数 17 页

[54] 发明名称 快速或慢速重放模式下不造成音调变化的音频信号重现方法及其重现装置

[57] 摘要

为了实施滤波处理,改变各个音频数据的呈现时间间隔以符合于用户指定的重放速度,使存储于存储器单元中解码的音频数据与改变的呈现时间间隔同步地以集合为单位写入一个输入队列中。对输入队列中的音频数据以帧为单位实施一种 TSM 算法,根据指定的重放速度,当指定的重放速度快于正常重放速度时减少音频数据数量,或者当指定的重放速度慢于正常重放速度时增加音频数据数量。



权 利 要 求 书

1. 一种响应于用户指定的一个重放速度值而重现出原来音频数据的方法，包含有步骤：

对于原来的音频数据按照一种时间标度调制算法实施时间标度调制处理，以响应于重放速度值而增加或减少原来音频数据的数量；以及

对于根据指定的重放速度值而由时间标度调制处理得到的音频数据，进行下取样或上取样，以使恢复出的取样的音频数据数量与原来的音频数据有同样的程序，

由此，当取样的音频数据以用户指定的重放速度重现时，取样的音频数据的音调实质上与原来音频数据的音调相同。

2. 权利要求 1 的音频数据重现方法，还包含有步骤，每当指令重放速度改变时，按照指定的重放速度值重新计算要增加/减少的音频数据的呈现时间间隔。

3. 权利要求 2 的音频数据重现方法，还包含有步骤，通过重新计算的呈现时间间隔重现出取样的音频数据。

4. 权利要求 1 的音频数据重现方法，其中，时间标度调制的步骤中包含有步骤：

将存储在一个缓存器工具中的原来的音频数据按每个预定的时间间隔以集合为单位写入一个输入队列中；以及

对存储在输入队列中的音频数据以帧为单位实施时间标度调制算法，当指定的重放速度快于正常重放速度时，根据指定的重放速度减少音频数据的数量，当指定的重放速度慢于正常重放速度时，根据指定的重放速度增加音频数据的数量，由此，将该音频数据提供给一个中间队列。

5. 权利要求 4 的音频数据重现方法，其中，取样步骤中包含有：

对于存储在中间队列中的音频数据，当指定的重放速度快于正常重放速度时，对之实施上取样处理，当指定的重放速度慢于正常重放速度时，对之实施下取样处理，其中，要传输至一个输出队列中的取样的音频数据的数量变为实质上与原来音频数据的数量相同；以及

将存储在输出队列中的取样的音频数据按每个预定的时间间隔以集合为单位传输至缓存器工具中。

6. 权利要求 5 的音频数据重现方法, 其中, 通过控制一种队列指针, 使输入队列、中间队列和输出队列之每一个的工作成为一个环形队列。

5 7. 权利要求 5 的音频数据重现方法, 其中, 将输出队列中取样的音频数据覆盖写入到缓存器工具中, 以替代缓存器工具中已存在的原来的音频数据。

8. 权利要求 4 的音频数据重现方法, 其中, 呈现时间间隔是一种重新计算出的取样音频数据的呈现时间间隔, 它是根据指定的重放速度值而变化的。

10 9. 权利要求 4 的音频数据重现方法, 其中, 写入到输入队列中的原来音频信号的集合数目是累加地计数的, 当计数的集合数目变到等于一帧中集合的数目时, 默认值为不启动状态的计算下一帧标志符便转移到启动状态, 由此以帧为单位实施时间标度调制算法。

15 10. 权利要求 4 的音频数据重现方法, 其中, 在单声道系统的场合下集合单元由一个音频数据构成, 在立体声系统的场合下集合单元由左/右声道的两个音频数据构成。

20 11. 权利要求 1 的音频数据重现方法, 其中, 在上取样/下取样中, 按照指定的重放速度值计算出数据数量的一个变化比, 由时间标度调制处理得到的音频数据的数量根据该变化比而变更, 但上取样/下取样之前和之后音频数据的特性实质上借助于应用一种内插方法而保持相同。

12. 权利要求 1 的音频数据重现方法, 其中, 根据指定的重放速度值由时间标度调制算法增加或减少原来的音频数据的数量, 同时保持着原来的音频数据的特性。

25 13. 响应于用户指定的重放速度, 在将经过 MPEG 系统中解码而存储在存储器装置中的解码音频数据供给至音频输出装置之前, 一种重现出解码音频数据的方法, 包含有步骤:

30 计算出指定的重放速度与正常重放速度之间的重放速度控制比, 将正常重放速度场合下解码音频数据的呈现时间间隔乘上该重放速度控制比, 产生出音频数据的一个新呈现时间间隔;

将存储在存储器装置中的解码音频数据以集合为单位写入一个输入队列中;

01-08-20

对于写入在输入队列中的音频数据,以帧为单位实施一种时间标度调制算法,以便与重放速度控制比成比例地增加或减少解码音频数据的数量,然后,将经过时间标度调制处理后的音频数据写入一个中间队列中;

- 5 对于写入在中间队列中的音频数据,在重放速度控制比小于 1 的快速重放模式场合下实施一种上取样,在重放速度控制比大于 1 的慢速重放模式场合下实施一种下取样,取样方式中,所应用的取样速率为重放速度控制比的倒数,以做到实施取样后的音频数据的数量实质上与解码音频数据的数量相同,再将取样音频数据传输至一个输出队列中;

10 将存储在输出队列中的音频数据以集合为单位写入至该存储器装置中,取代掉已存在的解码音频数据;以及

借助于所产生的呈现时间间隔重现出新写入存储器装置中的音频数据,

- 15 由此,即使指定的重放速度快于或慢于正常重放速度,重现声音的音调实质上与正常重放速度的音调相同。

14. 权利要求 13 的音频数据重现方法,其中,输入队列、中间队列和输出队列之每一个的工作,通过对其控制一个队列指针而如同一个环形队列。

- 20 15. 如权利要求 13 的音频数据重现方法,其中,集合的单位在单声道系统中由一个音频数据组成,或是在立体声系统中由左/右声道的两个音频数据组成。

16. 权利要求 13 的音频数据重现方法,其中,在上取样/下取样中,各个音频数据的幅度通过对于存储在中间队列中音频数据的幅度并根据重放速度控制比所实施的一种内插方法来确定,以使取样之前和之后的音频特性实质上保持相同。

17. 权利要求 13 的音频数据重现方法,其中,时间标度调制算法根据指定的重放速度值而增加或减少解码音频数据的数量,以保持解码音频数据的特性。

- 30 18. 一种根据用户指定的重放速度值对受到滤波处理后的音频数据进行重现的方法,包含有步骤:

响应于指定的重放速度值,增加或减少正常重放速度中音频数据的

呈现时间，并保持住音频数据的呈现时间间隔具有正常重放速度的值；

应用对于音频数据的一种预定的时间标度调制算法实施时间标度调制处理，以根据指定的重放速度增加或减少音频数据的数量；以及

5 在改变的呈现时间期内以该呈现时间间隔重现出从时间标度调制处理中得到的音频数据，

由此，即使指定的重放速度快于或慢于正常重放速度，重现声音的音调实质上与正常重放速度的音调相同。

10 19. 权利要求 18 的音频数据重现方法，其中，预定的时间标度调制算法根据指定的重放速度值而增加或减少解码音频数据的数量，以保持解码音频数据的特性。

20. 一种响应于用户指定的重放速度值对音频数据进行重现的装置，包含有：

15 重放速度控制装置，用于产生出指定的重放速度与正常重放速度之间的一个重放速度控制比，并将正常重放速度上音频数据的呈现时间间隔乘以该重放速度控制比，以产生出一个新的呈现时间间隔；

存储器装置，用于以包为单位存储规定的音频数据；

20 滤波装置，用于对存储在存储器装置中的音频数据根据预定的时间标度调制算法实施时间标度调制处理，以便根据指定的重放速度值增加或减少音频数据的数量，并对于从时间标度调制处理中得到的音频数据根据指定的重放速度实施下取样或上取样，以将取样音频数据的数量恢复至与时间标度调制处理之前音频数据的数量实质上相同的程度，并将取样的音频数据写入至存储器装置中以取代那里已存在的音频数据；以及

25 音频输出装置，用于以新的呈现时间间隔接收从存储器装置上来的滤波的音频数据，并将滤波的音频数据重现成声音。

由此，即使指定的重放速度无论快于或慢于正常重放速度，以新的呈现时间间隔重现时重现声音的音调实质上与正常重放速度的音调相同。

30 21. 权利要求 20 的音频数据重现装置，其中，预定的时间标度调制算法根据指定的重放速度值而增加或减少音频数据的数量，以保持音频数据的特性。

010500

22. 权利要求 20 的音频信号重现装置, 其中, 在上取样/下取样中, 滤波装置根据指定的重放速度值计算数据数量的变化比, 并根据该变化比对通过时间标度调制处理后得到的音频数据的数量进行变更, 而保持住音频数据在应用一种内插方法进行上取样/下取样之前和之后它们的特性实质上相同。

23. 一种重现音频信号的装置, 包含有:

音频信号供给装置, 用以响应于用户指定的重放速度值从记录媒体中读出并提供音频信号; 以及

数字信号处理装置, 它具有一个背景部分和一个前景部分, 背景部分一方面将音频信号供给装置来的音频数据以集合为单位写入输入队列; 另一方面又在同样的一个周期上以集合为单位读出存储在输出队列中的音频数据, 并将输出队列中读出的音频数据变换成模拟信号, 前景部分对于存储在输入队列中的音频数据以帧为单位应用一种预定的时间标度调制算法实施预定的时间标度调制, 以便响应于指定的重放速度值通过对时间标度调制处理后得到的音频数据根据指定的重放速度值实施下取样或上取样, 使取样音频数据的数量恢复至与时间标度调制之前音频数据的数量实质上相同的程度, 并将该取样的音频数据传输至输出队列中。

24. 权利要求 23 的音频数据重现装置, 其中, 数字信号处理装置中还包含有模拟/数字变换装置, 用于当从音频信号处理装置所加的音频信号为模拟信号时, 将音频信号供给装置与输入队列之间的模拟音频信号变换成数字数据。

25. 权利要求 23 的音频数据重现装置, 其中, 预定的时间标度调制算法根据指定的重放速度值而增加或减少音频数据的数量, 以保持音频数据的特性。

26. 权利要求 23 的音频数据重现装置, 其中, 在上取样/下取样中, 数字信号处理装置根据指定的重放速度值计算出数据数量的变化比, 并根据该变化比对通过时间标度调制处理后得到的音频数据的数量进行变更, 以保持音频数据在应用一种内插方法进行上取样/下取样之前和之后它们的特性实质上相同。

说明书

快速或慢速重放模式下不造成音调变化的 音频信号重现方法及其重现装置

5 本发明涉及一种重现音频信号或者音频/视频信号的方法，以及其重现装置。更具体地，涉及一种音频信号处理方法，在非正常重放速度的高速度或低速度下重现音频信号或者音频/视频信号时，它能够重现出不发生可感知音调变化的音频信号。

10 视频和音频节目信号按照一种确立的算法或方法变换成数字格式，并进行压缩、编码和复用。压缩的数字系统信号中，也即比特流中，包括有视频部分、音频部分和其它信息部分。这样的数据通过一种传输线路传输至一个重现装置上，或者存储入一个记录媒体中。安装有多媒体播放器能用以对复用的视频数据和音频数据中得到的多媒体数据进行重现的数字重现装置，诸如数字通用光盘（DVD）系统、
15 数字盒式录像机（VCR）或者计算机系统中，配置有解码装置，以重现上述的比特流信息。这种解码装置按照压缩算法实现去复用、去压缩并解码比特流，由此提供出可重现的信号。解码的视频和音频信号输出至诸如是荧光屏或扬声器之类的重现装置上，使信息呈现给用户供视听。

20 对视频和音频信号压缩和编码时，由合适的编码器加以实施，它执行一种选定的数据压缩算法，适应于在数字视频数据的发射机和接收机之间一致认可的标准或规范。高效的压缩标准已经由 MPEG（活动图像专家组）开发成功，包括 MPEG-1 和 MPEG-2，它们又被继续改善而出现了 MPEG-4。各种 MPEG 标准除了能在 VCR、DVD
25 或者类似的多媒体记录/重现装置中支持正常重放模式外，还能支持正向或反向的高速或低速重现。

各种 MPEG 标准中，根据一个称作标准目标解码器（STD）的理想解码器，约束了一种建议的同步方案。编码形式下的视频或音频数据单元或数据帧称为存取单元（AU），而未编码或已解码形式下的则
30 称为呈现单元（PU）。理想解码器中，视频和音频数据呈现单元取自基本流（ES）缓存器，在合适的呈现时间上即时地呈现给用户供视听。对一个呈现单元指明正确呈现时间的一种呈现时间标记（PTS），作为

系统句法的一部分在 MPEG 包头中传输。

呈现时间标记和存取单元不必需一起传输，因为它们由分层结构中的不同层次运载。所以，解码器必需与包层中找到的呈现时间标记相联系，第一个存取单元就在该呈现时间标记之后。实际情况更为复杂，在实际解码器中，系统对呈现单元的呈现时间有小量的控制。例如，视频解码器中，视频帧（图像）必须以帧频的准确倍数呈现，使视频图像显得匀滑，而音频帧必须以声音帧频的准确倍数呈现，使声音中没有喀喇声。

理想的 MPEG 同步方案中，在解码器里提供一个维持系统时钟时间的系统时间时钟（STC）。系统时钟时间的初始值在系统流内传输时，在 MPEG-1 比特流中由编码器作为系统时钟基准（SCR）进行传输，在 MPEG-2 比特流中由编码器作为节目时钟基准（PCR）进行传输。解码器将其本地系统时间时钟设定到该初始值上，然后，在 90 KHz 的时钟频率上使初始值不断增量。

随后，编码器对音频或视频存取单元传输一个呈现时间标记，过一些时间之后跟随有 AU 本身。解码器将 PTS 与本地系统时钟时间进行比较。如果它们相等，则从基本流缓存器中取出 AU 单元，立即解码，产生出相应的 PU 以使其呈现。

然而，在安装有多媒体播放器的常规多媒体重现装置诸如是 DVD 系统、数字 VCR 或计算机系统中，当用户选择快速或慢速重放模式时，视频数据按照指定的重放模式进行重现，但音频数据由于难以与视频数据保持同步而是静音抑制的。有所改善的重现装置，它们在快速或慢速重放模式中可以同时重现音频数据。然而，在此场合下，音频数据样本的呈现时间间隔在输出端只是根据指定的重放速度简单地增大或减小。更具体地，在快速重放模式下，有关音频数据的呈现时间间隔比之正常重放时的将变得缩窄，由于音频数据率被信频程而提升，使重现声音的音调升高；反之，在慢速重放模式下，有关音频数据的呈现时间间隔比之正常重放时的将变得加宽，由于音频数据受到信频程作用而丢失，使重现声音的音调降低。这样，导致所谓的音调变化现象。

上述的音调变化现象在诸如是模拟信号处理装置的 VCR 或盒式磁带记录器之类的视频记录/重现系统中，有着同样的表现。这类系统

中，如果用户将正常重放速度变更为高速度或低速度，则由重现装置从记录媒体中读出信号的速度将与变更的重放速度相关联，变高或变低。因此，当读出的音频信号不作改变地输出时，重现声音所给出的音调将变更，比之于正常速度下的重现声音会给出变高或变低的音调。

图 1 示出与一个 MPEG 重现装置之解码相关联的一种功能框图，用于重现 MPEG 文件。由一个文件源供给的 MPEG 文件通过数据输入器 10 后，通过数据分离器 12 分离成视频数据和音频数据。分离的视频数据和音频数据分别由视频解码器 14 和音频解码器 18 接收，依靠它们的解码而恢复出原来的数据，然后，它们分别提供给视频输出器 16 和音频输出器 20，以重现图像和声音。

附带地说，如果用户发出高速或低速重现的指令，则解码器 18 按照指定的重放速度之快速状态或慢速状态，改变音频包之包头中包含的 PTS 值。于是，有关音频样本的呈现时间间隔比之正常重放模式的将会压缩或扩展。当常规的 MPEG 文件处于快速或慢速重放模式时，重现出有关音频样本的时间间隔与正常重放的相比将受到压缩或扩展，使重现声音的音调听起来有变化，变更成为高音调或低音调。

音调变化之发生是由于常规重现系统在快速或慢速重现模式中时，只是在时间标度上简单地扩展或压缩有关音频信号的呈现时间间隔。更麻烦的是，为了防止音调变化，对快慢速度要独立地实施某种其它的信号处理。换句话说，在快速或慢速重现模式中，还需要一种附加的方案来防止音调变化。

考虑到先有技术中上面列举的问题，本发明的一个目的是提供一种重现方法以及相应的重现装置，它们对音频数据应用一种滤波处理，能够在将重放速度变更为高速或低速模式的场合下使重现的音频信号或者活动图像中包容的音频信号，在音调上与正常重放模式中的音调实质上相同。

为了实现本发明的上面的目的，按照本发明的一个方面，提供出了一种方法，它响应于用户指定的重放速度的快速状态或慢速状态，通过对音频数据进行滤波后重现音频数据。在使重现的音频数据进行滤波的方法中，它对于按照一种预定的时间标度调制算法中的音频数据，实施一种时间标度调制，即响应于所指定之重放速度的快速状态

或慢速状态，增加或是减少音频数据中的数据数量。随后，针对已按照指定之重放速度的快速状态或慢速状态而通过时间标度调制所得到的音频数据，实施一种下取样或上取样。当实现这种取样后得到的数据与解码的音频数据达到几乎相同的程度时，便恢复出了音频数据的数量。

借助于这种方法，在滤波之后实施音频数据重现，可使得尽管指定的重放速度慢于或快于正常重放速度，仍能维持重现声音的音调与正常重放速度的音调实质上相同。按照本重现方法，每当指令重放速度改变时，便根据指定之重放速度的加快或减慢而对音频数据的呈现时间间隔重新进行计算。又，借助于该新产生的呈现时间间隔重现出经实施取样处理后的音频数据。

更具体地，时间标度调制的步骤中包括有：将存储在缓存器装置中的原来的音频数据在每个预定的时间间隔上以一个集合为单位写入至一个输入队列中；当指定的重放速度快于正常重放速度时，根据输入队列中存储的音频数据以帧为单位实施时间标度调制算法，以按照指定的重放速度减少音频数据的数量；以及当指定的重放速度慢于正常重放速度时，便按照指定的重放速度增加音频数据的数量。由此，其音频数据给出至一个中间队列中。

更详细地说，取样步骤中包括有：对于中间队列内存储的音频数据，当指定的重放速度快于正常重放速度时，实施上取样处理，当指定的重放速度慢于正常重放速度时，实施下取样处理；于是，要传输给一个输出队列的已取样音频数据的数量将变得与原来的音频数据的数量实质上相同；随后，将存储在输出队列中的已取样音频数据在每个预定的时间间隔上以集合为单位写入至缓存器装置中。

这里，可取地，该预定的时间间隔是新计算出的取样音频数据的呈现时间间隔，它是根据指定的重放速度值而变化的。

更可取地，通过控制一个有关的队列指针，输入队列，中间队列和输出队列它们的工作如同一个环形队列。

输出队列的取样音频数据覆盖写入缓存器单元中，以取代缓存器单元中已存在的原来的音频数据，由此，使得实际重现的音频数据是已经按照本发明受到了滤波处理的音频数据。

如果用户指令其重现速度减慢到正常重放速度的 α 倍 ($\alpha > 1$)，则

响应于此指令，在对正常重放速度的呈现时间间隔进行比较之后，将有关音频数据的呈现时间间隔设定为增大至正常的 α 倍。这样，与正常重放速度下的八度音相比较，重现声音的八度音将发生降低到大约 α 倍的效应。在此情况下，通过应用时间标度调制算法使音频数据受到时间标度调制。这样做之后有一种效应，由于音频数据的数量增加至 α 倍，而增大了 α 倍的呈现时间间隔不改变，所以总呈现时间增大 α 倍。接着，在受到 $1/\alpha$ 倍的时间标度调制处理以实现减少音频数据数量的下取样之后，在时间标度上达到了压缩音频数据的效果。于是，对于其呈现时间间隔增大 α 倍下已经受到下取样处理的音频数据之重现，重现声音的八度音将从原来降低到 α 倍的八度音上提高 α 倍，从而恢复出正常重放速度的音调。

与之相反，如果用户指令其重现加快到正常重放速度的 α 倍（ $\alpha > 1$ ），则响应于此指令，在对正常重放速度的呈现时间间隔进行比较之后，将有关音频数据的呈现时间间隔简单地设定为减小至正常的 α 倍。这样，与正常重放速度下的八度音相比较，重现声音的八度音将出现升高到大约 α 倍的效应。在此情况下，通过应用时间标度调制算法使音频数据受到时间标度调制。这样做之后有一种效应，由于音频数据的数量增加至 α 倍，而增大了 α 倍的呈现时间间隔不改变，所以总呈现时间减小 α 倍。接着，在受到 $1/\alpha$ 倍的时间标度调制处理以实现增加音频数据数量的上取样之后，在时间标度上达到了扩展音频数据的效果。于是，对于其呈现时间间隔减小 α 倍下已经受到上取样处理的音频数据之重现，重现声音的八度音将从原来降低 α 倍的八度音上提高 α 倍，从而恢复出正常重放速度的音调。

考虑一下按照本发明之重现方法的一个优选实施例，它可以用作一种重现出解码的音频数据的方法以响应于用户指定的重放速度，而后将 MPEG 系统中已经解码的、存储在存储器单元中的音频数据提供给音频输出单元。按照本发明的此实施例，计算出指定的重放速度与正常重放速度之间的一个重放速度控制比 α 。又，通过对正常重放速度下音频数据的呈现时间间隔乘上该重放速度控制比，再计算其音频数据之新的呈现时间间隔。将存储在存储器单元中的音频数据以集合为单位写入至输入队列中。相对于写入在输入队列内的音频数据，以帧为单位实施一种预定的时间标度调制算法，以做到与重放速度控制比 α

成比例地增加/减少音频数据的数量，并将实施时间标度调制算法中得到的已调制音频数据写入至中间队列中。此后，相对于写入在中间队列中的音频数据，在重放速度控制比 α 小于1的快速重放模式下，实施一种上取样，或是在重放速度控制比 α 大于1的慢速重放模式下，实施一种下取样。取样频率决定于重放速度控制比 α 的倒数，以使得实施取样处理后音频数据的数量实质上与解码的音频数据相同。将取样处理后的音频数据馈送至一个输出队列中。然后，存储在输出队列中的音频数据以集合为单元装载至存储器单元中，并使该存储器单元内的音频数据由计算出的呈现时间间隔作出重现。采用这种方法，即使指定的重放速度快于或慢于正常重放速度，重现声音的音调仍实质上与正常重放速度的相同。

这里，在受到下取样或上取样处理后传输给输出队列的有关音频数据之数量的确定，是借助于相对于存储在中间队列内音频数据的数量，根据重放速度控制比 α 的值实行一种内插方法来做到的，以使得实质上相同地保持住在取样之前和之后的音频数据的特性。另外，在音频数据的数量按照指定之重放速度的加快或减慢而增加/减少时，任何算法可以实施作为预定的时间标度调制算法，同时几乎能保持住执行滤波之前音频数据的全部特性。

与此同时，这里提供出一种装置，用于与上面的重现音频数据方法相符合的音频信号重现。按照本发明的一个实施例，该音频信号重现装置能响应于用户指定的一种重放速度而重现出音频数据。这里，该重现装置中包括一个重放速度控制单元，用于计算指定的重放速度与正常重放速度之间的一个重放速度控制比，并进一步通过对正常重放速度下的音频数据呈现时间间隔乘上重放速度控制比 α ，计算出一个新的呈现时间间隔。重现装置中包括一个存储器单元，用于以一种在数据包单元内可区辨的格式存储入数字音频数据。又，包括有一个滤波单元作为其一部分，它相对于存储在存储器单元中的音频数据根据一种预定的时间标度调制算法实施时间标度调制，响应于指定的重放速度而增加/减少音频数据的数量，相对于根据指定的重放速度通过时间标度调制而得到的已调制音频数据，实施下取样或上取样，并将取样后音频数据的数量恢复至实质上与解码的音频数据之数量相同的程度上。此外，重现装置中有一个音频输出单元，它接收新的呈现时间

间隔下存储单元来的、由音频数据滤波单元已作出处理的音频数据，并重现出声音。采用这种构造，无论用户指定的重放速度快于或慢于正常重放速度，由新的呈现时间间隔进行重现时重现声音的音调实质上保持与正常重放速度的音调相同。

5 按照本发明的另一个实施例，一个重现音频信号的装置中有一个供读出用的音频信号供给单元，以响应于用户指定的快速模式或慢速模式重放速度而从记录媒体中提供出音频信号，并有一个数字信号处理单元，它包含一个背景部分和一个前景部分。背景部分同时实施音频数据的写入和读出，即将音频信号供给单元的音频数据以集合为单位写入至输入队列中，又在同一时期上以集合为单位对输出队列内存储的音频数据进行读出，并将输出队列中读出的音频数据变换成模拟信号。另外，该数字信号处理单元的前景部分针对存储在输入队列内的音频数据，应用一种预定的时间标度调制算法以帧为单位实施一种预定的时间标度调制，以便按照指定的重放速度值增加或减少数据数量。前景部分还根据指定的重放速度而经由时间标度调制处理得到的音频数据，实施一种下取样或上取样，以便在执行取样处理之后将音频数据的数量恢复至实质上与解码的音频数据相同的程度，并将取样处理中得到的音频数据装载至输出队列中。可取地，该装置中还有一个模拟/数字变换单元，当供给自音频信号处理单元的音频信号为模拟信号时，用于将模拟音频信号变换成在音频信号供给单元与输入队列之间的数字数据。

10

15

20

又需指出，在音频数据的数量按照指定的重放速度值增加/减少时，任何算法可以实施作为预定的时间标度调制算法，同时几乎保持住执行滤波之前音频数据的全部特性。其时，更可取地，下取样或上取样处理的实施是通过根据指定的重放速度值而确定一个增加/减少比的，并在按照增加/减少比而对音频数据的数量进行增加/减少时，实质上保持了应用内插方法进行取样之前和之后同样的音频数据特性。

25

按照本发明的另一个方面，提供出一种方法，根据用户指定的重放速度值将受到滤波处理后的音频数据进行重现。本方法中包括有步骤：响应于指定的重放速度值，增加或减少正常重现速度中音频信号的呈现时间，并将音频数据的呈现时间间隔维持为具有正常重放速度时的值；针对音频数据应用一种预定的时间标度调制算法实施一种时

30

同标度调制处理，以便根据指定的重放速度值增加或减少音频数据的数量；以及在改变的呈现时间期内对时间标度调制处理中得到的音频数据以该呈现时间间隔进行重现。因此，即使指定的重放速度快于或慢于正常重放速度，重现声音的音调实质上与正常重放速度的音调相同。

按照本方法，有关音频数据的呈现时间和呈现时间间隔都适合于调整至使重现声音的音调与正常重放速度的音调相同，无论上取样/下取样处理怎样被删除。又，此种场合下，在音频数据的数量按照指定的重放速度之快速或慢速模式增加/减少时，任何算法可以实施作为预定的时间标度调制算法，同时几乎保持住执行滤波之前音频数据的全部特性。

参考附图，通过对优选实施例的详细叙述，本发明的上述目的和其它优点将变得十分显明，各附图中：

图 1 示出一个用于重现一种多媒体文件的功能方框图，涉及通常的 MPEG 重现装置中的解码；

图 2 示出一个涉及 MPEG 重现装置解码的功能方框图，用于实施一种按照本发明之方法以对音频信号的滤波进行处理；

图 3 是一个流程图，概略示明本方法的一个顺序，用于实施在一个实时的时间标度调制 (RTTSM) 滤波器中执行的音频信号滤波；

图 4、图 5 和图 6 各是流程图，分别示明 RTTSM-put 功能、RTTSM-out 功能和 RTTSM-calc 功能的程序；

图 7 示例出将缓存器内有关的音频包变换成新的音频包的过程，它是通过执行 RTTSM-put 功能、RTTSM-out 功能和 RTTSM-calc 功能经由输入队列 Q_x 、中间队列 Q_y 和输出队列 Q_z 而受到滤波处理的新音频包；

图 8 示例出一种得到一个时间标度修改信号 $y(\cdot)$ 的原理，使得相对于原来的音频信号 $x(\cdot)$ 执行该时间标度修改算法，以响应于用户设定的重放速度而增加或减少数据数量；

图 9 示明在低于正常重放速度 2 倍的指定慢速重现场合下，按照本发明的每一数据处理步骤中音频数据之呈现时间间隔的状态改变；

图 10 示明在高于正常重放速度 2 倍的指定快速重现场合下，按照本发明的每一数据处理步骤中音频数据之呈现时间间隔的状态改变；

图 11 示明按照本发明用于下取样处理的一种方法，又图 12 示明按照本发明用于上取样处理的一种方法；

图 13 示明按照本发明用于处理音频数据的数字信号处理板的组成，它附加到诸如是磁带录音机、VCR、DVCR 和 DVCR 系统之类的音频或音频/视频重现装置上，能在慢速或快速重放模式下使重现声音的音调象正常重放速度中那样地保持不改变；

图 14(a)是一个流程图，示明前景部分的数据处理程序，图 14(b)是一个流程图，示明背景部分的数据处理程序，又，图 14(c)示例出执行前前景部分和背景部分切换时的关系图；以及

图 15 是应用中断信号作为基准时关于背景部分和前景部分工作时间的定时图。

作为开始，将说明按照本发明之一方面重现出音频信号的一种方法以及其装置。

图 2 示出一个涉及 MPEG 重现装置解码的功能方框图，用于示例一种按照本发明对音频信号进行处理的方法。该 MPEG 重现装置与图 1 中所示常规装置的不同点在于，音频解码器 18 先使通过去压缩和解码后由压缩音频数据中恢复出的原来的音频包供给至 RTTSM 滤波器 22 上，然后将它供给至音频输出器 20 上，以按照本发明实施一种音频信号滤波处理。

经执行去压缩和解码后从音频解码器 18 上得到的输出数据暂时存储入数据包单元内的一个输出缓存器 24 中。这里假定，用户指定了重放速度为低速度重现（例如是慢 2 倍）或是高速度重现（例如是快 2 倍）。当与正常重放期间解码出的数据（对应于图 9(a)）相比较时，输出缓存器 24 上记录的音频数据将变成这样的数据（对应于图 9(b)），它响应于变化的重放速度而在时间标度上修改为具有分别修改好的呈现时间间隔。在此种运行中，MPEG 重现装置响应于用户指定的快速或慢速模式的重放速度，通过扩展或缩短方式实现一种处理，以重新设定呈现时间间隔。也就是，需要以一种方式来实现一项处理，它计算出用户指定的重放速度与正常重放速度之间的一个重放速度控制比 α ，并对正常重放速度的音频数据呈现时间间隔乘上重放速度控制比 α 以产生出一个新的音频数据呈现时间间隔。本发明提出的音频信号重现装置提供一种工具，也即一个程序，每当用户通过重现装置的一个

键钮输入单元（未示出）改变重放速度时，它响应于指定之重放速度的加快或者减慢，重新产生出有关音频数据的呈现时间间隔。并且，根据计算出的呈现时间间隔，重现出已受到按照本发明之滤波处理的音频数据。这样，由诸如是一个 CPU（未示出）的控制工具执行该提供供给重现装置的程序。这里，当指令低速度重现要比正常重放速度慢 1.5 倍时，重放速度控制比 α 的值变为 1.5，或者当指令高速度重现要比正常重放速度快 2 倍时，重放速度控制比 α 的值变为 0.5。换句话说，重放速度控制比 α 是由指定的重放速度与正常重放速度之间速度比的倒数决定的。

10 存储在缓存器 24 内的有关音频包（此后称之为原来的信号 $x(\cdot)$ ，它有时是指在处理 RTTSM 滤波之前的音频信号）按照本发明变为 RTTSM 滤波器 22 的输入数据（参见图 7）。

由 RTTSM 滤波器 22 实现的音频数据滤波处理概略地示明于图 3 的流程图中。RTTSM 滤波器 22 的功能可以用软件或者硬件的方式来实现。首先，参考图 3 的流程图说明 RTTSM 滤波器 22 的功能。

15 RTTSM 滤波器 22 实施的基本功能是响应于用户指定的快速或慢速重放速度，对输出缓存器 24 提供来的输入队列 Q_x 中音频数据之数据量作出增加/减少，这是音频数据的时间标度修改（TSM），并将它存储入中间队列 Q_y 内作为一个 TSM 信号 $y(\cdot)$ 。音频数据的 TSM 可以应用已知的诸 TSM 算法之一来实现，不需对之作任何特别的修改，或者为了符合于应用目标而对之做某些修改。

已经提出了几种音频信号处理技术，用于调整由用户所指定的音频信号重放速度。具体地，有几种已知的音频信号处理技术，它们能够在时间标度基础上以增加或减少数据数量的方式变更重放速度，而同时保持住与原来音频信号固有的那些特性相类同的特性。其中，由 Roucus 和 Wilgus 于 1985 年提出的一种重叠-相加（OLA）算法是一种代表性的技术。在它被引入之后，OLA 算法已经发展成同步 OLA（SOLA）和基于波形相似性的 OLA（WSOLA）等。此外，还已经知道有修正或改善 OLA 算法的一些技术，诸如整体和局部搜索的时间标度修改（GLS-TSM），时域音高同步 OLA（TD-PSOLA）和指针间隔控制 OLA（PICOLA）等。

下面对本发明的说明中采用的是利用 WSOLA 技术作为一种

RTTSM 算法。考虑 WSOLA 算法时，应用一个预定大小的窗口将音频数据切割成许多块，使得两个接连的块由一个规则的时间间隔搭叠上，然后，由对应于速度变化的时间间隔进行重新安排之后将各块相加，以使原来的信号变换成在时间标度上增加或减少了的数据。这样，WSOLA 算法可以产生出已变换的信号，它能够在不同于原来重放速度的一个速度上进行重现。然而，如果在改变时间标度间隔之后简单地使相互不同的块的信号相加，它们将导致声音质量有下降，很大地不同于原来信号的声音质量。当各块重新安排时，为了使时间标度修改的信号其声音质量最大地类似于原来信号的声音质量，需要有一种相关性估计，在某一个范围内提供出对时间标度间隔作小量调整；调整到所需的基本间隔，同时启动对两个信号之间波形相似性的判定。然后，通过移动两个块的信号将它们加以综合，直到时间标度间隔的小量调整对应的值具有最大的波形相似性。这样做之后，就可能无论怎样改变重放速度，声音质量总能保持在与原来声音的质量几乎相同的程度上。WSOLA 算法基于上面所述的概念。也就是，WSOLA 算法的特征在于，为了防止在重新安排而综合两个块的信号中声音质量的下降，将两个接连块的信号移动一个间隔，以便使两个接连块的两个重叠部分之间的波形相似性有一个最大值。

所以，本发明采用 WSOLA 算法进行说明，它已知是音频数据的几种 TSM 算法中使重现的原来声音具有优良特性的一种，例如对应于用户指定的重放速度控制比 α 的大小，对输入队列 Q_x 上积累的原来音频信号 $x(\cdot)$ 的长度进行增加或减少而得到 TSM 信号 $y(\cdot)$ 。然而，应当指出，本发明并非仅仅适用于 WSOLA 算法，也能适用于任何其它算法，只要它们结合前述的 TSM 算法可响应于重放速度而在时间标度基础上对音频数据的数据量一般地进行增加或减少。

在对适用于 WSOLA 算法的 RTTSM 滤波进行处理中，首先，在步骤 S10 中，每周期地检查用户是否指令了变化重放速度而需变更先前设定的重放速度。如果不指令变化重放速度，则按照已经设定的重放速度实施处理。如果指令了变化重放速度，则重现装置发生一个事件。

步骤 S12 上，根据事件的发生，重新规定环境参数以符合于所指定的重放速度。环境参数是实施 WSOLA 算法所需要的，它们包括：

0 1 0 5 20

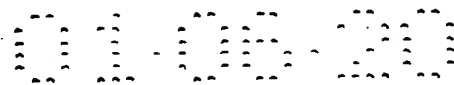
原来信号 $x(\cdot)$ 的帧尺寸 n ；在先前帧与当前帧之前具有一个最佳相关性的同步滞后 K_m ，由一个检验范围（窗）规定的 K_{max} ，用于按照原来信号 $x(\cdot)$ 与 TSM 信号 $y(\cdot)$ 之间的波形相似性来检测最佳相关性； S_x ，它定义为从缓存器 24 中读出的接连各帧原来信号 $x(\cdot)$ 的各开始时间之间诸音频样本的差值； S_y ，它定义为经 WSOLA 算法处理后得到的 TSM 信号 $y(\cdot)$ 之接连各时间标度帧的各开始时间之间诸音频样本的差值；TSM 信号 $y(\cdot)$ 的接连帧之间重叠样本 L 的数目；以及加权值 W_x 和 W_y ，它们施加到 TSM 信号 $y(\cdot)$ 的接连帧之间重叠的样本上。这里，参数 n 、 K_{max} 、 S_x 、 W_x 和 W_y 是在重现装置中先前定义了的，参数 S_y 是对当前重放速度乘上参数 S_x 得到的一个值，参数 L 是由一个相对公式 $L = n \cdot S_y$ 定义的值，而同步滞后 K_m 是指下一帧原来信号 $x(\cdot)$ 开始位置的最佳校准点，在这个最佳校准点上原来信号 $x(\cdot)$ 与 TSM 信号 $y(\cdot)$ 具有最佳的相关性。

此外，在步骤 S14 中，对一些参数进行再初始化，包括：环形队列 Q_x 、 Q_y 和 Q_z ，它们使用于实施算法中的暂时存储工具；队列指针；取样频率；一个标志符，它用于按照重现装置的性能合适地调整计算量；以及各个设定值，它们是检验波形相关性中规定出一些单元所必需的。

执行该算法后，响应于用户指定的新重放速度建立起了环境参数，针对先前存储于缓存器 24 中已经由音频解码器处理的已解码音频数据，RTTSM 滤波器 22 通过应用 WSOLA 算法对该指定的重放速度起响应而增加或减少数据数量。然后，数据再经过下取样或上取样并返回到缓存器 24 中。因此，供给至音频输出器 20 的数据是已经由 WSOLA 算法进行处理并受到下取样或上取样后的数据。

更具体地，存储在缓存器 24 中的音频数据重复地经受图 3 中步骤 S18 至步骤 S24 的处理，直至到达每个包的结束，并且相同的处理对于随后的音频包重复进行。

借助于实施三种功能，完成对于各个音频包的 RTTSM 滤波处理，这三种功能是 RTTSM-put 功能、RTTSM-calc 功能和 RTTSM-out 功能。在步骤 S18 中，RTTSM-put 功能从缓存器 24 中读出一个集合的音频数据（对应于图 9(b)），将它写入输入队列 Q_x 。RTTSM-calc 功能对帧单元内积累在输入队列 Q_x 上的音频数据实施 WSOLA 算法处



理，以响应于指定的重放速度而增加或减少数据数量。这样，通过对当前重放速度的响应，形成了数据数量已经增加或减少而时间标度改变的音频数据 $y(\cdot)$ （对应于图 9(c)），写入至中间队列 Q_y 中。在步骤 S20 上，当当前指定的重放速度慢于正常重放速度时，积累在中间队列 Q_y 上的音频数据被下取样，以减少数据数量，或是当当前指定的重放速度快于正常重放速度时，积累在中间队列 Q_y 上的音频数据被上取样，以增加数据数量，并将取样的数据写入至输出队列 Q_z 中。又，在步骤 S22 上，RTTSM-out 功能再将积累在输出队列 Q_z 中的音频数据向缓存器 24 提供给一个集合，由此，以实施 RTTSM 滤波处理后得到的音频数据取代由音频解码器 18 供给的已有的音频数据。

步骤 S24 上，每当顺序地完成三种功能之后，使音频集合的指数增大 1，然后返回至步骤 S16。在步骤 S16 上，重复上述程序，直至音频数据集合的指数值变得与音频包最后样本的指数相同，并对随后的音频包重复上面说明的数据处理程序。先前存储在缓存器 24 中的原来的音频包按照上面的程序变化成一个新的音频包，以此种方式变化成的音频包在重现装置的控制下输出至音频输出器 20 中，由此通过扬声器（未示出）或是耳机端子（未示出）重现出声音。

通过 RTTSM 算法处理得到的新音频包由音频输出器 20 予以重现，具有的音调与正常重放速度的音调实质上相同，不依赖于用户指定的重放速度。得到如此结果的原由将参照图 4 至图 10 予以说明。

图 9 中示出，当指定的重放速度比之正常重放速度慢 2 倍时，每个有关的数据处理步骤中音频数据呈现时间间隔的变化。图 9(a)示明对应于正常重放速度的音频数据呈现时间间隔。假定在正常重放速度期间各个音频数据 $d1$ 、 $d2$ 、……、 $d10$ 、……的呈现时间间隔为 t ，图 9(b)中示出音频解码器 18 产生的数据，各个音频数据 $d1$ 、 $d2$ 、……、 $d10$ 、……的呈现时间间隔简单地增大至 2 倍，并使产生的音频数据存储入缓存器 24 中。由于存储入缓存器 24 中的各个音频数据 $d1$ 、 $d2$ 、……、 $d10$ 、……的呈现时间间隔为 $2t$ ，所以音频数据的重现时间也扩展至 2 倍。如果音频数据的呈现时间间隔在时间标度上增大至 2 倍，尽管能够满足用户所需的重放速度，但重现声音的音调将降低大约一个八度音，结果是劣化了重现声音的质量。

存储在缓存器 24 中的音频数据如图 9(b)中所示地预定以 $2t$ 的时间

间隔呈现。为了按用户的指令那样以慢 2 倍的重放速度重现它，并维持与视频数据的同步、且能使重现声音的音调与正常重放速度下原来的音调相同，首要地对缓存器 24 中存储的音频数据实施 WSOLA 处理。如图 9(c) 中所示，通过实施 WSOLA 处理得到的音频数据的数据量增加至大约 2 倍。那时，由于各个音频数据的呈现时间间隔仍维持 $2t$ 那样大，将不发生音调变化。

另外，在用户指定了 2 倍的低速度重现的场合下，重现装置将音频数据的呈现时间间隔从 t 改变到 $2t$ 。结果，受到 WSOLA 处理后的音频数据其重现时间增加为大约 4 倍，它牵涉到这样的问题，与视频数据的同步变得不协调，并且音调仍然保持为降低一个八度音的音调变化。

为了解决这些问题，使实施 WSOLA 算法后得到的音频数据受到下取样。对于实施下取样，从概念上假定，相对于实施 WSOLA 算法之后得到的音频数据，如图 9(d) 中所示地将音频数据的呈现时间间隔在时间标度上压缩至要恢复成的 t 。在实现此种处理之后，总的重现时间变为如图 9(b) 中所示。因此，音频数据的重现能够符合于用户设定的新重放速度，并可以与视频数据同步。此外，由于时间标度上有着再压缩 $1/2$ 的效应，音频数据的音调被提升一个八度音地恢复，而几乎与图 9(a) 中所示的音调相同。

然而，与图 9(b) 中所示的相比较，音频数据样本的数目仍保持为它的 2 倍，且重现装置预安排好呈现出每 $2t$ 的音频数据。由于这些情况，仅仅在时间标度上的压缩是不够的。换句话说，对于按照 $2t$ 的呈现时间间隔来重现音频数据，需要使用 9(c) 中所示经 WSOLA 算法处理的音频数据其数量减少一半。为此，针对经 WSOLA 算法处理的音频数据实施下取样，以使音频数据的数量减少一半。通过实施下取样，最后得到如图 9(e) 中所示的音频数据。

由于图 9(e) 中所示的音频数据是对音频数据（对应于图 9(d)）进行下取样得到的，后者是图 9(c) 的音频数据在时间标度上压缩一半而音调升高一个八度音的，所以图 9(e) 的音调仍然与图 9(d) 音频数据的音调相同，从而与图 9(a) 音频数据的音调相同。结果，虽然重放速度减慢 2 倍，重现声音的音调保持为几乎与正常重放速度的音调相同。当然，在实施下取样的同时音频数据的分辨率降低了，但是，在实施下取样

期间施加上后面要说明的一种声音质量减低方法时，由分辨率降低造成的声音质量劣化是可以忽略的。

同时，图 10 中示出了当指定的重放速度比之正常重放速度快 2 倍时，每个有关的数据处理步骤中音频数据呈现时间间隔的变化。图 10(a) 5 示明实施正常重放速度期间音频数据 S1、S2、……、S10、……的呈现时间间隔。当用户指令 2 倍的快速重放时，重现装置使各个音频数据的样本呈现时间间隔压缩至 $1/2$ ，也即从 t 变为 $t/2$ ，如图 10(b) 中所示。存储在缓存器 24 中的音频数据在按照它那样重现时要以 $\Delta t/2$ 的时间间隔进行重现。因此，与正常重放速度的音调相比较，重现声音的 10 音调升高一个八度音。所以，音频数据要以这样的方式进行处理，即针对缓存器 24 中存储的音频数据执行 WSOLA 处理和上取样，以便不仅使重放速度加快至 2 倍，而且在重现声音中保持正常重放速度的音调。

首先，存储在缓存器 24 中的音频数据受到 WSOLA 处理，如图 15 9(c) 中所示地使音频数据的数量实质上减少 $1/2$ 。此时，由于各个音频数据的呈现时间间隔继续保持为 $t/2$ 不变，所以与正常重放速度的音调相比较，其音调亦保持为升高一个八度音的状态。实施 WSOLA 处理后音频数据的重现时间与正常重放速度的重现时间相比较缩短到 $1/4$ ，引起了与视频数据同步不协调的问题，并关联到音调变化保持于 20 高出一个八度音的问题。

为了解决这些问题，对于实施 WSOLA 处理后得到的音频数据进行上取样。关于实施上取样，从概念上假定，相对于自 WSOLA 处理 25 中得到的音频数据，如图 10(d) 中所示地将音频数据的呈现时间间隔在时间基线上扩展 2 倍而成为要恢复的 t 。实现这样的处理之后，总重现时间变为如图 10(b) 中所示。因此，音频数据的重现能够符合于用户设定的新重放速度，并可以与视频数据同步。此外，由于时间标度上有着再扩展 $1/2$ 的效应，音频数据的音调被提升一个八度音地恢复，而与图 10(a) 中所示的实质上相同。

然而，与图 10(b) 中所示的相比较，音频数据样本的数目仍保持为 30 它的 $1/2$ ，且重现装置预先安排好呈现出每 $t/2$ 一个的音频数据。由于这些情况，只是在时间标度上压缩是不足够的。换句话说，对于按照 $t/2$ 的呈现时间间隔来重现音频数据，需要使用经实施 WSOLA 处理得到

的音频数据，其数据数量增加 2 倍。为此，对于自 WSOLA 处理中得到的音频数据实施上取样，以使其音频数据的数量增加 2 倍。通过实施上取样，最后得到如图 10(e) 中所示的音频数据。

由于图 10(e) 中所示的音频数据 S_1 、 S_2 、……、 S_{10} 、……是对音频数据（对应于图 10(d)）进行上取样得到的，后者是图 10(c) 的音频数据在时间标度上扩展 2 倍而音调降低一个八度音的，所以图 10(e) 的音调仍然与图 10(d) 音频数据的音调相同，从而与图 10(a) 音频数据的音调相同。结果，虽然重放速度加快到 2 倍，重现声音的音调保持为几乎与正常重放速度的音调相同。

执行 WSOLA 算法后，上述的下取样或上取样由后面将说明的三种功能来实现。另外，在实现下取样或上取样中，数据的增加率或减少率是根据用户指定之重放速度的加快或减慢状态确定的，音频数据之数量的增加或减少就按照所确定的增加率/减少率。经取样后各个音频数据的幅度可以照原样地采用自 WSOLA 处理中得到的 TSM 音频数据的幅度，或者可以由相邻音频数据中内插出的幅度来确定。下面，将说明应用各别功能进行的特定的数据处理算法。

图 4、图 5 和图 6 的流程图分别示明 RTTSM-out 功能、RTTSM-out 功能和 RTTSM-calc 功能的程序。图 7 示例出将缓存器 24 中有关的音频包变换成新音频包的过程，它借助于实施三种功能而历经了输入队列 Q_x 、中间队列 Q_y 和输出队列 Q_z 。图 8 示例出得到 TSM 信号 $y(\cdot)$ 的原理，使得响应于用户设定的重放速度的加快或减慢状态，原来的音频信号 $x(\cdot)$ 的长度也即音频数据的数量，在时间标度上被扩展或压缩。本发明中，利用了三个队列以应用三种功能来实施 WSOLA 处理和上取样/下取样。

当用户指定一个重放速度比之正常重放速度加快或减慢时，重现装置便改变各个音频数据集合的呈现时间间隔以适合于所指定的重放速度。所以，存储在缓存器 24 中的各个音频包集合由图 9(b) 或图 10(b) 中所示的呈现时间间隔预定地重现。缓存器 24 中的每个音频包在音频数据的集合指数增加 1 时顺序地由一个集合 d_{xj} 予以读出，并如图 4 中所示地通过实施 RTTSM-out 功能而在步骤 S30 中写入至输入队列 Q_x 中。这里，集合有两种组成类型。对于单声道系统，它可取地由一路音频数据组成；对于立体声系统，它就能由两路音频数据组成，用于

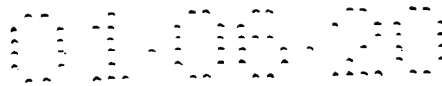
左/右声道。毋庸说明，在音频数据的阶序保持不变的前提下，一个集合中可以包含多路音频数据。

输入队列 Q_x 可取地要求其大小足以在它上面能积累大约 3 帧以上的音频数据。当写入一个集合时，输入队列 Q_x 的一个指针值便增大。在增大队列指针的过程中队列指针指明是输入队列 Q_x 的最后位置之后，便复位而指明开始位置，使输入队列 Q_x 起到一个环形队列的作用。此外，当输入队列 Q_x 上写入一个集合后，它被计数。然后，步骤 S34 中，当计数的集合数目变为与参数 S_n 的集合值相同时，一个计算下一帧的标志符改变成启动状态，用以确定是否对下一帧进行计算。当然，该计算下一帧标志符的默认值为 1，表示为不启动状态，而标志符改变为 0 值的启动状态时，表明输入队列 Q_x 中存储了至少一帧，能实施 WSOLA 算法。

在按照本发明实施滤波之前，与输入队列 Q_x 上写入音频数据的同时，从缓存器 24 中读出一个集合，如图 5 中所示的 RTTSM-out 功能在其实现中读出存储在输出队列 Q_x 内已经受到 WSOLA 处理和上取样/下取样处理的音频数据之一集合 d_{ij} ，然后在步骤 S36 中，当集合指数增大 1 时用输入场合时同样的速率将该集合覆盖写入缓存器 24 中。由于实施 WSOLA 处理和下取样/上取样处理后的数据数量与实施处理前的相同，所以除了总的重现时间延时一短段时间（也即是实施 WSOLA 处理和下取样/上取样处理所需的时间）之外，不发生别的问题，即使是从输出队列 Q_x 中读出一个集合而顺序地写入缓存器 24 中也是那样。输出队列 Q_x 的大小设定为能同时存储入至少两帧的数据，并在步骤 S38 上调整队列指针用以起到环形队列的作用。

在将积累在输入队列 Q_x 上的音频数据传输至输出队列 Q_x 期间，执行图 6 中所示的 RTTSM-calc 功能以实施基于 WSOLA 算法的 TSM 处理和下取样/上取样处理。应当指出，虽然 RTTSM-put 功能和 RTTSM-out 功能的执行周期是以集合为单位的，但 RTTSM-calc 功能的处理以帧为单位，它是多个集合构成的组群。也就是，只当步骤 S40 中计算下一帧的标志符为启动状态时，才实施 RTTSM-calc 功能。另外，每当对于当前帧实现上述的处理时，在步骤 S42 中计算下一帧标志符的值转移至不启动状态，以准备处理下一帧。

在由 RTTSM-out 功能处理当前帧期间，检查是否有改变重放速



度的指令。这样，当存在改变重放速度的指令时，在步骤 S44 和 S46 上，再修改在处理 WSOLA 算法中所需的环境参数的值，也即重放速度、 S_s 、 S_r 、 L 、 W_x 和 W_y 的值。在处理 WSOLA 算法期间，由重现系统作出控制，产生一个事件以使得 RTTSM-calc 功能每次处理 S_s 单位的音频数据。此时，由于必需按照同一个重放速度来处理一个 S_s ，因重放速度改变引起的修改环境参数的程序固有地包括在 RTTSM-calc 功能的程序内。

当重放速度不改变时，如下面那样以预先设定的环境参数值来实施 WSOLA 处理。在执行 RTTSM-put 功能中，输入队列 Q_x 用音频数据积累起来。这里，当每计算下一帧的标志符设置于启动状态一次时，便对存储在输入队列 Q_x 中的音频数据实施 RTTSM 处理。为了实施 WSOLA 处理，需要输入队列 Q_x 中至少存储一帧音频数据。

为了有助于理解 WSOLA 处理，下面将叙述示例的解释。假定，帧 F_m 由 320 个样本组成，重放速度控制比 α 为 2，在输入队列 Q_x 上原来信号 $x(\cdot)$ 的接连帧之间开始点的基本差值 S_s 为 120 (样本)，以及在原来信号 $x(\cdot)$ 与 TSM 信号 $y(\cdot)$ 之间检测波形相似性之最佳相关度的检验范围(窗口) K_{max} 为 ± 40 (样本)。重放速度控制比 α 表示 WSOLA 处理期间数据数量的变动。例如，如果指定的重放速度比之正常重放速度慢 2 倍，则重现装置将重放速度控制比 α 设定为 2，使得与处理前的数量相比较，WSOLA 处理后音频数据的数量增大至大约 2 倍。由于数据数量增大至 2 倍，重现时间将增大 2 倍以使得重放速度减慢 2 倍。与之相反，当重放速度指定为比之正常重放速度快 2 倍的高速重放时，重放装置便将重放速度控制比 α 设定为 0.5。这样，与处理前的数量相比较，WSOLA 处理后音频数据的数量减少至大约 1/2 倍。重现时间也将象数据数量的减少那样缩短那么多。另外，假定音频数据是对于单声道系统产生的，则其中一个样本构成一个集合。下面的说明也可应用于立体声系统，因为从基本的数据处理系统来看，除了在立体声系统的场合下一个集合由两个样本组成这一事实要给予任一种特定的考虑外，它与单声道系统并无其它差别。上面这些值只是给出作为例子，按照其作用的环境它们可以改变为不同的值。

开始时，原来信号 $x(\cdot)$ 的初始帧 F_0 从输入队列 Q_x 中读出，以便复写入中间队列 Q_y 作为 TSM 信号 $y(\cdot)$ ，并将帧指数 m 的值设置为 1。

此后, 从下一帧起执行下面的环路, 直至整个原来的音频信号 $x(\cdot)$ 完全地修改好。

然后, 从输入队列 Q_x 中读出原来信号 $x(\cdot)$ 的下一帧 F_1 , 以作为 TSM 信号 $y(\cdot)$ 加到中间队列 Q_y 中。这里, 当从原来信号 $x(\cdot)$ 中读出一帧 F_1 时, 读出开始位置可能不一样地确定。该读出开始位置由对于先前读出帧 F_0 的同步滞后 K_1 确定, 帧 F_0 是先前复写入中间队列 Q_y 作为 TSM 信号 $y(\cdot)$ 的。另外, 将作为 TSM 信号 $y(\cdot)$ 的、读出的下一帧 F_1 加到中间队列 Q_y 中的哪个位置, 是根据重放速度控制比 α 的大小而可变化地确定的。

同步滞后 K_m 应用来寻找出一个点, 在这个点上原来信号 $x(\cdot)$ 的下一帧 F_m 与已经作为 TSM 信号 $y(\cdot)$ 记录在中间队列 Q_y 中的先前帧 F_{m-1} 在一个预定的范围具有最佳的相关性。

下面的相关性公式应用来确定同步滞后 K_m 。同步滞后 K_m 的计算是通过使 F_m 围绕在最小值 $ms - 40$ 和最大值 $S_s + 40$ 之间的范围 $y(m \cdot S_s)$ 内滑动进行的。

$$c(m, k) = \sum_{j=0}^{L-1} y(mS_s + j) \cdot x(mS_a + k + j), S_s = \alpha S_a \dots \dots (1)$$

$$C(m, K_m) = \max [c(m, k)], -K_{max} \leq K \leq K_{max} \dots \dots (2)$$

式中, L 是接连帧中重叠样本的数目。

这里, 在计算接连帧之间的最佳相关性时, 可应用使音频数据逐一地滑动的计算方法。然而, 这种计算方法增加了重现系统在实施许多计算上的负担。所以, 当要求加快计算速度时, 一种跳越多个音频数据的方法可以推荐作为最佳相关性的计算方法。然而, 不可避免地, 从最佳相关性的精确度看, 该方法劣于前者的方法。在决定采用哪种方法时, 可取地应考虑到重现装置中 CPU 的性能更适合于哪一种。

通过应用上面两个公式得到了具有最佳相关性的同步滞后 K_m 之后, 借助于应用所得到的同步滞后 K_m , 从原来信号 $x(\cdot)$ 中读出包括有 N 个样本数目的下一帧 F_m 。该下一帧 F_m 的读出开始位置是这样一个位置, 它决定于从先前帧 F_{m-1} 的读出开始位置上跳越 $S_s \pm K_m$ 个样本数目, 这里, $0 < S_s \pm K_m < N$ 。例如, 参见图 8, 如果 K_1 、 K_2 和 K_3 分别被决定为 20、-10 和 35, 则是第二、第三和第四帧 F_1 、 F_2 和 F_3 的

读出开始位置分别是第 140、第 230 和第 395 个样本。当然，各个帧中样本的数目 N 恒定地为 320。按照上面说明的从输入队列 Q_x 中读出原来信号 $x(\cdot)$ 的方法，前一帧和下一帧中相当数量的样本是重叠的。另外，根据重放速度控制比 α 的大小，TSM 信号 $y(\cdot)$ 中的重叠部分可以是大的 ($\alpha > 1$ 时) 或是小的 ($\alpha < 1$ 时)。应当指出，各个帧的读出开始位置不是按 S_a 与帧指数 m 的乘积规则地改变的，而是按照应用公式 (1) 和 (2) 所决定的最佳相关性 K_m 的大小作不规则地变化的。

如上面所述的读出之后，将下一帧 F_m 加到中间队列 Q_y 中的 TSM 信号 $y(\cdot)$ 上。要加到 TSM 信号 $y(\cdot)$ 上的下一帧 F_m 的各个开始位置决定于 $m\alpha S_a$ 。所以，当图 8 中 $\alpha S_a = 2 \times 120 = 240$ 时，对于相加到第二、第三和第四帧 F_1 、 F_2 和 F_3 上的开始位置分别变为 240、480 和 720。该相加中，在下一帧 F_m 的前部与前一帧 F_{m-1} 的后部之间有着重叠部分。接连帧的各重叠部分是加权地相加的，加权值应用下面的公式 (3) 和 (4) 得出，而下一帧 F_m 的其余部分只是简单地复写。

$$Y(mS_a+j)=(1-g(i))y(mS_a+j)+g(j)(mS_a+K_m+j), 0 \leq j \leq L \cdots \cdots (3)$$

$$Y(mS_a+j)=(mS_a+K_m+j), L \leq j \leq N \cdots \cdots (4)$$

这里， $g(j)$ 是一个加权值函数，它的代表形式可取地为一个线性函数。另一种情况，也可以采用指数函数作为加权值函数。

上面所述的状态下，存储在缓存器 24 内、呈现时间间隔如图 9(b) 或图 10(b) 中所示的原来信号 $x(\cdot)$ 以帧为单位从输入队列 Q_x 中读出，然后相加到中间队列 Q_y 的 TSM 信号 $y(\cdot)$ 上。通过此处理，中间队列 Q_y 上积累了响应于指定的重放速度，已经受到 WSOLA 处理的音频数据。例如，如果指定的重放速度比正常重放速度慢 2 倍，则将图 9(c) 中所示的音频数据存入中间队列 Q_y 中，如果指定的重放速度比正常重放速度快 2 倍，则将图 10(c) 中所示的音频数据存入中间队列 Q_y 中。对于中间队列 Q_y ，可取地确保其大小能存储至少两帧音频数据。

然后，积累在中间队列 Q_y 中经过 WSOLA 处理的音频数据传输至输出队列 Q_z 上。在传输期间，按照重放速度实施下取样或上取样。实施取样中，根据用户指定的重放速度确定一个数据增加率/减少率，然后，按照确定的增加率/减少率应用一种内插方法使音频数据的数量改变，能使取样之前和之后的数据特性不致发生任何变化。内插方法是一种数值分析方法，用于从诸个给定的点上推断出一个新的点。有几

种典型的内插方法：一种应用 Taylor（泰勒）多项式的内插方法，它一般使用于数值解释中；一种应用 Lagrange（拉格朗日）多项式的内插方法；一种重复内插方法；一种 Hermite 内插方法和三维 Spline 内插方法；以及一种最为简单的线性内插方法。任一种内插方法都可适用于本发明，只要它能够使音频数据的特性在取样之前和之后几乎相互一样。

后面将假定，借助线性内插方法来实施内插，它应用了通过所给定全部音频数据点的一个公式和一些加权值。

当用户指定比正常重放速度为慢的重放速度时，在数据的数量已经由 WSOLA 处理而增加后，应当实施下取样处理。将参考图 11 予以说明。

作为图 11(a)中所示的第一个例子的情况，指定的重放速度比正常重放速度慢 2 倍，也即重放速度控制比 α 设定为 2。存储在中间队列 Q_y 中、受到 WSOLA 处理后的音频数据 $d1'$ 、 $d2'$ 、 $d3'$ 、……的数量在传输至输出队列 Q_z 之前要先减少一半。为了减少音频数据的数量，每两个接连的音频数据或者音频样本应减少为一个音频数据。此时，新形成的一个音频数据的量值也即幅度，是通过对两个接连的音频数据幅度进行平均而得到的一个值。也就是，实施下取样后音频数据 $d1''$ 、 $d2''$ 、 $d3''$ 、……的幅度值有如下关系式：

$$d1''=(d1'+d2')/2$$

$$d2''=(d3'+d4')/2$$

$$d3''=(d5'+d6')/2$$

.....

图 11(b)示明第二个例子的情况，指定的重放速度比之正常重放速度慢 1.5 倍，也就是重放速度控制比 α 设定为 1.5。存储在中间队列 Q_y 中、受到 WSOLA 处理后的音频数据 $d1'$ 、 $d2'$ 、 $d3'$ 、……的数量在传输至输出队列 Q_z 之前要先减少到 $2/3$ 倍。在此场合下，下取样以这样的方式实施，即看作为一个单元的三个音频数据减少至为两个音频数据，每三个音频数据中第一个音频数据不作变化，对第二和第三个数据计算出平均值并将它们减少成为一个音频数据。这样，从下取样中得到的音频数据 $d1''$ 、 $d2''$ 、 $d3''$ 、 $d4''$ ……的幅度值具有下面的关系式：

$$d1''=d1'$$

$$d2''=(d2'+d3')/2$$

$$d3''=d4'$$

$$d4''=(d5'+d6')/2$$

.....

5 图 11(c) 示明第三个例子的情况, 指定的重放速度比之正常重放速度慢 1.25 倍, 也即是重放速度控制比 α 设定为 1.25。存储在中间队列 Q_y 中, 受到 WSOLA 处理的音频数据 $d1'$ 、 $d2'$ 、 $d3'$ 、..... 的数量减少至 4/5 倍。下取样以这样的方式实施, 即看作为一个单元的五个音频数据减少至为四个音频数据。也就是, 要传输至输出队列 Q_z 的、从下

10 取样中得到的音频数据 $d1''$ 、 $d2''$ 、 $d3''$ 、 $d4''$ 、 $d5''$ 、..... 的幅度值具有下面的关系式:

$$d1''=d1'$$

$$d2''=d2'+(d3'-d2')/4$$

$$d3''=d3'+2(d4'-d3')/4$$

$$d4''=d4'+3(d5'-d4')/4$$

$$d5''=d6'$$

.....

可以知道, 应将上面的方法修改成规一化形式, 将它应用到重放速度控制比 α 为任何值的场合。

20 另一方面, 当用户指定了快速重放模式时, 对于存储在中间队列 Q_y 中的音频数据应实施上取样, 在此场合下, 音频数据由 WSOLA 处理予以减少后, 通过上取样形成的新音频数据传输至输出队列 Q_z 中。将参考图 12 详细进行说明。

25 图 12(a) 示明第一个例子的情况, 指定的重放速度比之正常重放速度快 4/3 倍, 重放速度控制比 α 设定为 0.75。在此场合下, 存储在中间队列 Q_y 中的音频数据 $d1'$ 、 $d2'$ 、 $d3'$ 、..... 的数量在传输至输出队列 Q_z 之前需要由上取样增加至 4/3 倍。也就是, 从音频数据 $d1'$ 、 $d2'$ 、 $d3'$ 、..... 经上取样 4/3 倍得到的音频数据 $d1''$ 、 $d2''$ 、 $d3''$ 、 $d4''$ 、 $d5''$ 、 $d6''$ 、..... 的幅度值具有下面的关系式:

$$d1''=d1'$$

$$d2''=d1'+3(d2'-d1')/4$$

$$d3''=d2'+2(d3'-d2')/4$$

$$d4''=d3'+(d4'-d3')/4$$

$$d5''=d4'$$

$$d6''=d4'+3(d5'-d4')/4$$

.....

5 图 12(b) 示明第二个例子的情况, 指定的重放速度比之正常重放速度快 2 倍, 也就是, 重放速度控制比 α 设定为 0.5。在此场合下, 存储在中间队列 Q_y 中经过实施 WSOLA 处理后音频数据 $d1'$ 、 $d2'$ 、 $d3'$ 、..... 的数量要增加至 2 倍。对此, 将每两个相邻音频数据的平均值采用来形成一个新音频数据, 它随后插入在这两个相邻的音频数据之间。这样, 从上取样中得到的音频数据 $d1''$ 、 $d2''$ 、 $d3''$ 、 $d4''$ 、 $d5''$ 、..... 的幅度值类似地有下面的关系式:

$$d1''=d1'$$

$$d2''=(d1'+d2')/2$$

$$d3''=d2'$$

$$15 \quad d4''=(d2'+d3')/2$$

$$d5''=d3'$$

.....

按照上面说明的下取样或上取样, 存储在中间队列 Q_y 中的每个音频数据帧的数据数量根据所指定重放速度的快速或慢速状态增加或减少, 并且在传输至输出队列 Q_z 上之前, 上取样/下取样的音频数据的幅度值在确定中或是采用存储在中间队列 Q_y 中的音频数据之幅度值, 或是根据上面的规则采用对它们进行内插得到的幅度值。应当指出, 上面的说明中各个重放速度控制比 α 的值是示例性的, 当指定任何其它不同的重放速度时, 上面的下取样和上取样的规则也是一般地可予应用的。

25 通过下取样或上取样得到的新音频数据以帧为单位传输至输出队列 Q_z 上。在实施 RTTSM-out 功能后, 输出队列 Q_z 中的音频数据以一个个集合为单位顺序地写入缓存器 24 中。这样做之后, 缓存器 24 中已有的音频包由已经受到 WSOLA 处理和下取样/上取样而从输出队列 Q_z 来的相应的新音频包取代。要提供给音频输出器 20 的音频数据即是这新的相应的音频包。

按照本发明所提出的重现装置, 每当指令重放速度改变时, 响应

于指定的重放速度值，要重新计算音频数据呈现时间间隔的增加/减少，而音频输出器 20 的工作情况是由新计算出的呈现时间间隔来重现经取样得到的音频数据。所以，即使指定的重放速度不同于正常重放速度时，重现声音的音调实质上与正常重放速度中重现的声音音调相同，并且在与视频数据的同步上不发生问题。

本发明引入了输入队列，中间队列和输出队列三种数据存储工具，用于 TSM 处理和上取样/下取样处理。但可以理解到，不需要对它们以物理方式来区分，重现装置中的一个存储器可以划分成三个区域而加以合适地利用。此外，定义出三个队列是为了便于软件的实现，其实不需要象上面那样地定义三个独立的队列。换句话说，可以用另一种方式定义这些队列，一个统一的全规模队列划分成三个区域，三个区域之每一个通过控制它的一个指针规定它起一个环形队列的作用。

如上面所述的按照本发明的音频数据处理方法，可以用软件方法实施而直接应用于计算机上，计算机内装置 Windows 操作系统和一个微软公司的称为 Direct Media 的程序。在软件实现方法中，实现音频数据处理方法的算法程序存储在计算机内的硬盘（未示出）中或是 ROM 240 中，并当实现一个多媒体重现程序时由 CPU 230 来执行。缓存器 24 或是三个环形队列 Q_x 、 Q_y 和 Q_z 适当地利用计算机内的 RAM（未示出）资源，并利用计算机内的一个声卡（未示出）作为音频输出器 20。

应用本发明之音频数据处理方法的可能性并不限制于计算机。本方法也能够应用于 DVD 系统 100a、数字 VCR 系统中或是另外的类似的系统中，也即应用于重现压缩编码之视频数据和音频数据的任何数字重现装置中。而且，可以应用于模拟系统的磁带录音机、VCR 系统 100 b 中，或是类似的系统中。换句话说，按照本发明的处理音频数据的方法能广泛地应用于无论是模拟系统或是数字系统，与压缩方法或编码方法也无关，而只要重现装置是涉及音频数据处理的。这样，就模拟系统的重现装置而言，是使音频信号变换成数字信号，实施按照本发明的 RTTSM 滤波处理，并将它再变换成模拟信号予以重现。

这类重现装置可响应于用户指定的重放速度值，从记录媒体中读出音频信号。

图 13 示出按照本发明的一种用以处理音频数据的数字信号处理 (DSP) 板 200 的构成, 它可以使用于模拟基重现装置内有代表性的诸如 VCR 系统的音频/视频重现装置中, 或是数字基重现装置内有代表性的 DVD 系统中, 并能够维持重现声音的音调而不受指定的重放速度变化的影响。

毋庸置疑, 重现装置中配备有一个重放速度控制部分, 用于计算用户指定的重放速度与正常重放速度之间的重放速度控制比 α , 以及计算正常重放模式的音频数据呈现时间间隔乘上重放速度控制比 α 后的新呈现时间间隔。由按钮输入 (未示出) 和诸如是微计算机的控制器和 CPU 230 等组合成的构体, 可以起重放速度控制部分的作用。

从硬件基础的角度看, DSP 板 200 可以由下列各部分构成: ROM 240; RAM (未示出), 并通过规定 RAM 资源可在其中确保三个队列; CPU 230 或 DSP 芯片; 振荡器 (未示出); 模/数变换器 (ADC) 210; 数/模变换器 (DAC) 220; 以及其它等。实现 RTTSM-calc 功能的程序安装在 ROM 240 中, RAM 的工作可以实现输入队列 Q_x' 、中间队列 Q_y' 和输出队列 Q_z' 。对 ADC 210 供给以伺服系统 100 来的、记录在录像带上的音频信号, 使它变换成数字数据。DAC 220 将数字数据变换成模拟信号, 使它能通过扬声器 300 重现出声音。CPU 230 顺序地执行存储在 ROM 240 中的安装程序, 以实施几种数据处理, 用于将 ADC 210 的输出数据写入输入队列 Q_x' 中, 将积累在输出队列 Q_z' 上的音频数据传输给 DAC 220, 并且对积累在输入队列 Q_x' 上的音频数据实施上述的 RTTSM-calc 功能后得到的音频数据进行 WSOLA 处理和下取样/上取样。当记录在记录媒体上的源信号象模拟 VCR 中那样记录成模拟信号时, ADC 210 是必需的。但当源信号为 DVD 系统中那样的数字信号时, 就不需要 ADC 210。

DSP 板 200 由一个背景部分 200a 和一个前景部分 200b 组成。背景部分 200a 在硬件基础上实现音频数据处理的功能, 将 ADC 210 的输出数据写入到输入队列 Q_x' 中, 并将积累在输出队列 Q_z' 上的音频数据传输给 DAC 220。前景部分 200b 实现的功能是将实施 WSOLA 处理和下取样/上取样、并随后针对输入队列 Q_x' 中存储的音频数据按照程序执行 RTTSM-calc 功能后得到的数据, 传输至输出队列 Q_z' 中。也就是, 背景部分 200a 在硬件基础上起到前述的 RTTSM-put 功能

和 RTTSM-out 功能的作用。换句话说,背景部分 200a 同时实现写入操作和读出操作,前者是将音频信号供给部分 100 a 或 100 b 中的音频数据以集合为单位写入输入队列 Q_x' 中,后者是以集合为单位读出存储在输出队列 Q_z' 中的音频数据,并将输出队列 Q_z' 中读出的音频数据
5 变换成模拟信号。前景部分 200b 的作用是针对存储在输入队列 Q_x' 中的音频数据应用一种诸如 WSOLA 的预定的 TSM 算法,以帧为单位实施 TSM 处理,以响应于指定之重放速度的快速或慢速状态而对数据数量进行增加/减少,并对于根据指定的重放速度而通过 TSM 处理得到的音频数据实施下取样/上取样,以使受到取样后音频数据的数量程度
10 恢复至与传输至输出队列 Q_z' 上的原来的音频数据数量实质上相同。

图 14(a)是一个流程图,示明前景部分 200 b 上的数据处理程序;图 14(b)是一个流程图,示明背景部分 200a 上的数据处理程序;图 14(c)示明在前景部分 200 b 与背景部分 200a 之间的切换关系。图 15 是在应用中
15 断信号作为基准下关于背景部分 200 a 和前景部分 200 b 工作时间的定时图。

在前景部分 200 b 执行诸如是 TSM 处理和上取样/下取样处理工作的同时,背景部分 200 a 对用于实现此类处理的音频数据的输入/输出进行控制。结果,前者要求的处理时间长于后者要求的处理时间。因此,希望 CPU 资源这样地工作,即每当供给中断信号时 CPU 230 将给出
20 出足以供执行背景部分 200a 业务用的资源,同时由前景部分 200 b 实现 RTTSM-calc 功能,并当中断服务程序 (ISR) 完毕时,就在执行上述的 ISR 之前通过返回到前景处理程序而继续进行前景部分 200 b 的处理。

参考图 14(a)和图 14(b),将更详细地说明执行前景部分 200 b 和背景部分 200a 的程序。一旦开始重现工作,DSP 板 200 上的 CPU 230
25 顺序地读出存储在 ROM 240 中的程序,以实现重现出音频数据的处理。首先,在步骤 S60 中,使输入队列 Q_x' 、输出队列 Q_z' 和实施 WSOLA 处理的各个环境参数初始化。

当通过对重现装置中一个振荡器所提供的时钟信号计数而每个恒定周期上产生出中断信号时,有着默认值为不启动的一个 ISR 的值转移至启动上,在步骤 S64 至 S72 内,每当 ISR 被启动时实现背景部分
30 200a 的数据处理。由于前景部分 200b 针对实现背景部分 200a 的 ISR

所得到的音频数据实施滤波处理，所以在步骤 S74 上执行一个无限循环的程序，直至下一帧开始标志符转移入启动状态。

为了实施 ISR 处理，步骤 S64 上，CPU 230 从 ADC 210 中取出一个集合的音频数据，并从诸如是按钮输入（未示出）的用户界面上分别地取出由用户指定的重放速度。步骤 S66 上，将 ADC 210 来的音频数据写入到输入队列 Q_1' 中。在步骤 S68 和 S70 上，在写入至输入队列 Q_1' 中一个集合时积累地计数一个值，检查一个计数值是否达到了包括在单个帧内的总集合数目。如果是这样，使初始设定于不启动状态的下一帧开始标志符的值转移入启动状态。上面的处理等效于前述的 RTTSM-put 功能的处理。差别在于，ADC 210 的输出数据写入在输入队列 Q_1' 中。接着，步骤 S72 上，CPU 230 访问输出队列 Q_2' ，以读出存储在那里一个集合的音频数据，将它传输至 DAC 220 中。这种处理等效于 RTTSM-out 功能的处理。只当如图 15 (b) 中所示背景脉冲维持高电平时，才实现上面那样的 ISR 处理。

与此同时，前景处理一旦起始后，设计成执行一个无限循环的程序。更详细地说，如果下一帧开始标志符的值设置为启动状态，则在步骤 S76 上使下一帧开始标志符的值转移成它的基本设定值即不启动状态。此后，在步骤 S78 上，按照前面的方法对于存储在输入队列 Q_1' 中的音频数据执行 RTTSM-calc 功能。然后，将处理好的音频数据传输至输出队列 Q_2' 中，并停留在那里直至输出到 DAC 220 上。

当用户指定重放速度为加快或减慢模式而不是正常模式时，通常的 VCR 系统以指定的重放速度从磁带上读出数据，它就像通常的重现装置从诸如是磁带录音机、DVCR 和 DVD 之类的记录媒体上读出数据予以重现一样。结果，供给至 ADC 210 的模拟音频信号是在时间标度上已压缩或扩展而其音调有变更的数据，ADC 210 的输出信号只是改变成了数字数据，仍保持着已改变的音调状态原封不动。因此，ADC 210 的输出数据可认为等效于在按照本发明实施 RTTSM 滤波处理之前存储在缓存器 24 中的数据。这样，在接收到这种数据以实现 RTTSM-calc 功能之后，形成的音频信号所保持的音调实质上与正常重放期间的音调相同，并且在与视频信号的同步上不发生不协调性，其结果达到了本发明的目的。

另一方面，当应用于数字 VCR 系统时，整体的数据处理系统几乎

与前面的情况一样，微小的差别只是在于，由于原来的信号是数字信号，所以 DSP 板 200 中不需要 ADC 210。类似地，在 DVD 系统中，尽管存在着记录媒体是 DVD 而不是磁带的差异，但由于事实上原来的信号仍是数字信号，所以 DSP 板 200 的组成中也不需使用 ADC 210，
5 总的数据处理情况几乎与前面的情况相同。

以上，按照本发明的一个方面，已经给出了如下情况的说明，即按照指定的重放速度，通过对系统中各个音频数据的呈现时间间隔值施加扩展/压缩作用来重现音频数据。按照上面的方法，由于应当对应于指定的呈现时间间隔来重现和输出音频数据，所以对音频数据进行
10 下取样或上取样的处理是需要的。

然而，按照本发明的另一个方面，按照指定之重放速度的快速或慢速状态，对音频输出器 20 进行控制，以扩展/压缩音频数据的总体呈现时间，同时保持各个音频数据的呈现时间间隔象正常重放速度的数值一样。按照这个特性，在慢速重放模式或者快速重放模式的场合下，
15 不需要下取样或上取样。更具体地是这样控制的，以正常重放速度作为基准，响应于指定的重放速度值使音频数据集合的总体呈现时间扩展/压缩，而音频数据的呈现时间间隔保持为正常重放速度的值。与此同时，通过应用上面说明的 TSM 算法对于音频数据实施 TSM 处理，以根据用户指定的重放速度值增加/减少数据数量。然后，在改变的呈现
20 时间期内控制受到 TSM 处理的音频数据，以其呈现时间间隔进行重现。对于重现该音频信号的信号处理使用上述的前面那种方式予以实施之后，重现声音的音调又保持与正常重放速度的音调实质上相同，不受指定的重放速度值的影响。它们优点在于，可以取消音频数据的取样，而能使声音质量较接近于原来的声音。

按照上面所述的本发明，即使是用户指定的重放速度是快速或慢速模式而不是正常重放速度，也可以这样处理音频数据，通过用改变的呈现时间间隔进行重现，能使重现声音的音调与正常重放速度的音调实质上相同。结果，它可以广泛地应用于诸如是模拟 VCR 系统、数字 VCR 系统、计算机系统和 DVD 系统之类的多媒体重现装置上，以
30 所希望的重放速度同时重现音频数据和视频数据，它给出的音调具有良好的质量。

此外，按照本发明的音频数据处理方法，可以独立于视频数据的

处理予以实施。所以，它能广泛应用于上面提到的几种多媒体重现装置中。换句话说，将其中配置有按照本发明对音频数据实现滤波处理之方法的模块，简单地相加到各个媒体重现装置的音频信号处理模块上，由此便能形成这样的媒体重现装置，它具有按照本发明音频数据重现功能。

虽然，参照具体实施例已经详细地示明和叙述了本发明，但本技术领域内的熟练人员知道，对此可在形式和细节上作出各种改变，它们偏离不开所附的权利要求书中确定的本发明的精神和范畴。

说明书附图

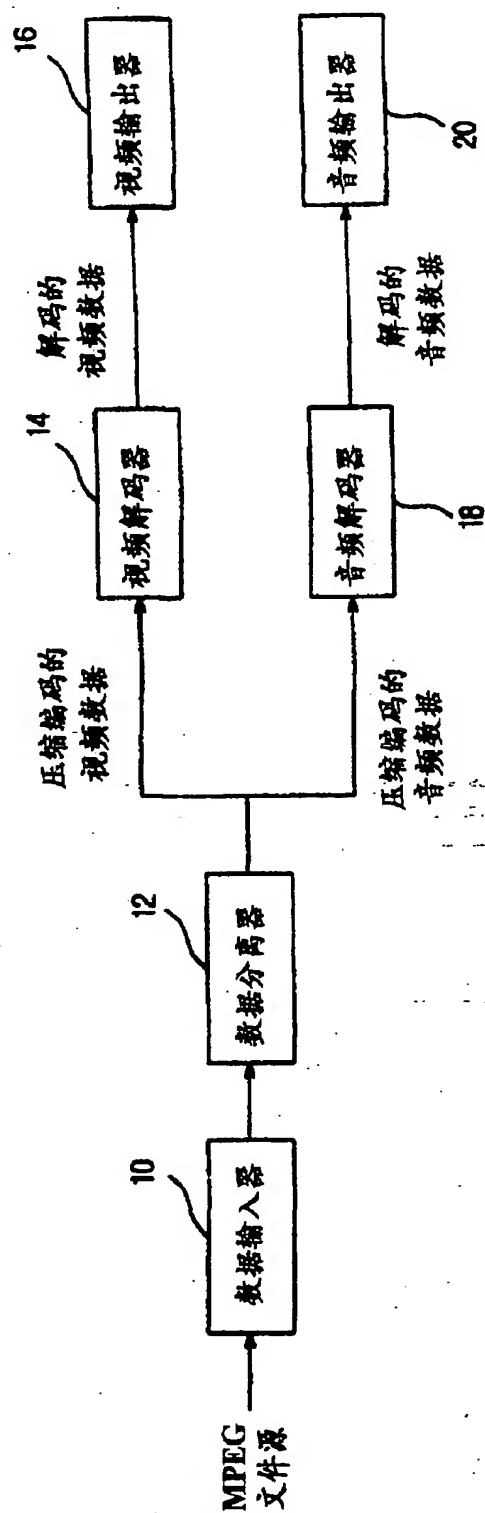


图 1

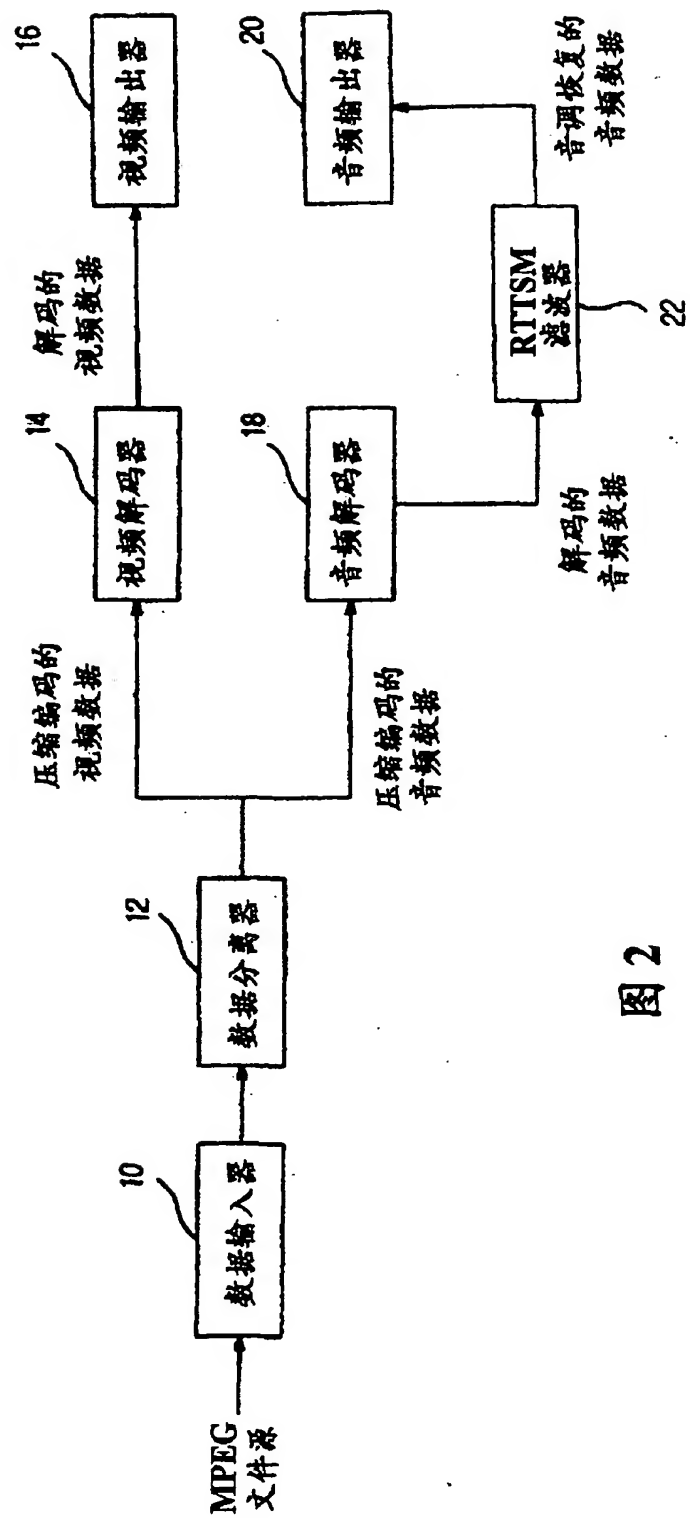


图 2

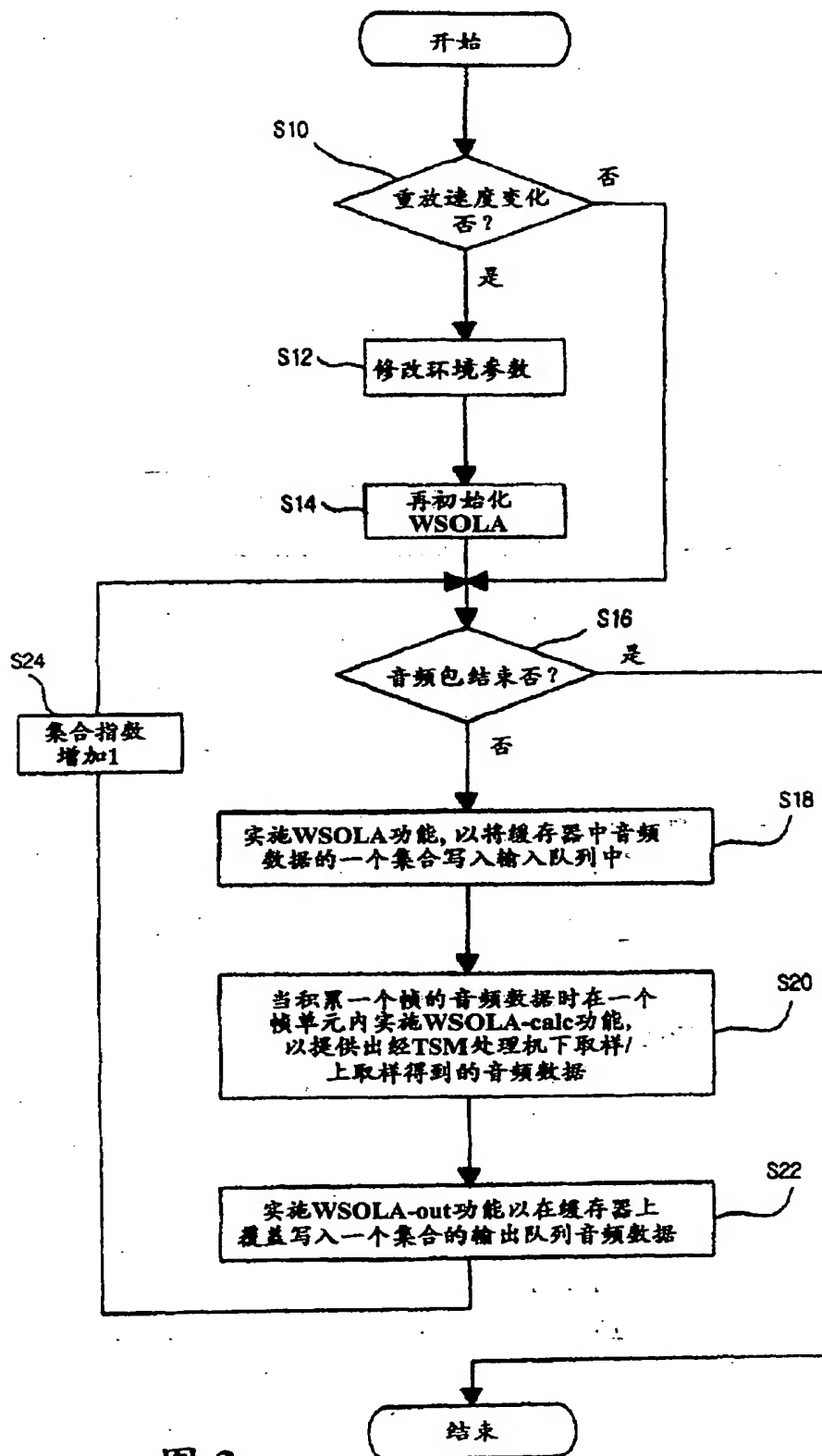


图 3

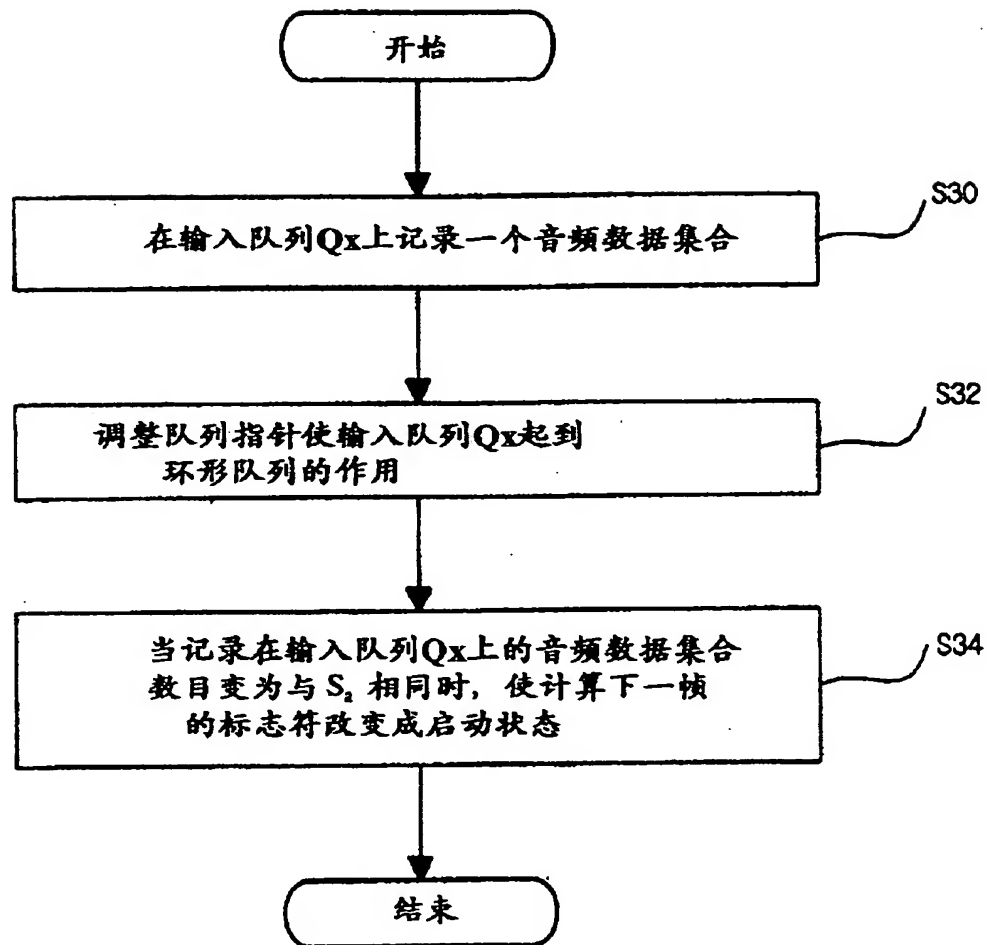


图 4

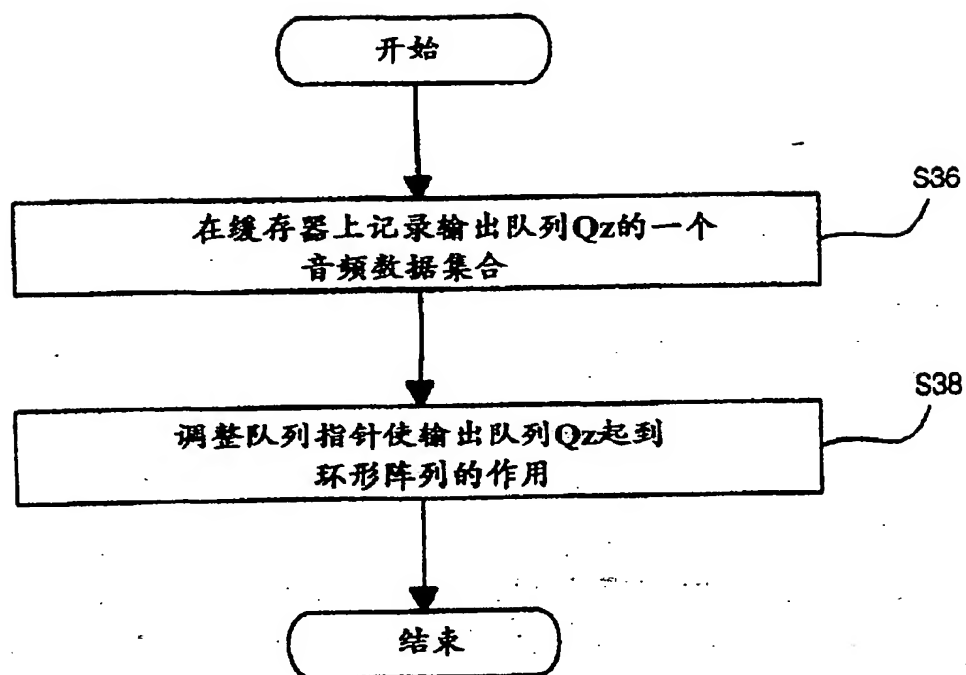


图 5

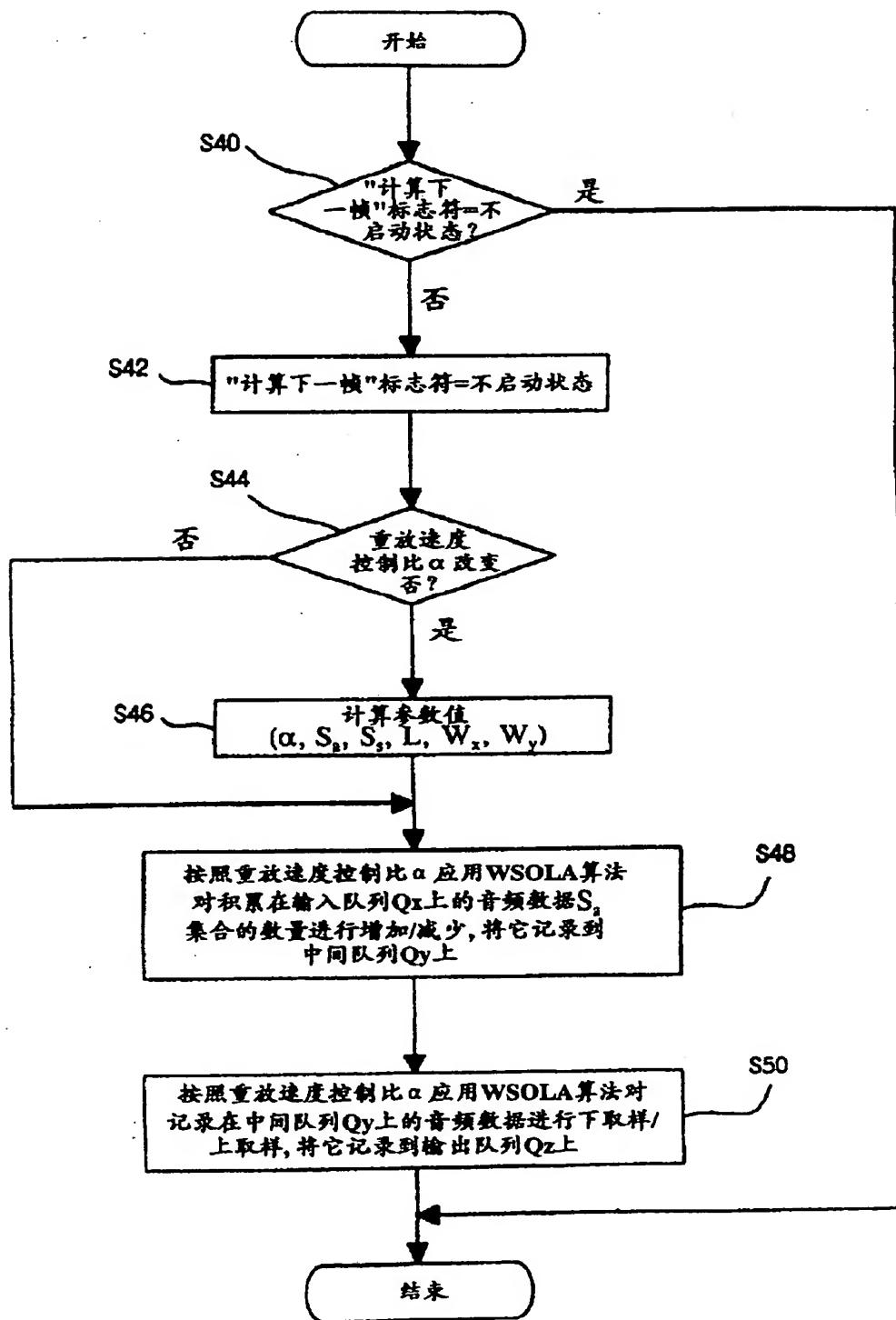


图 6

缓存器 (24)

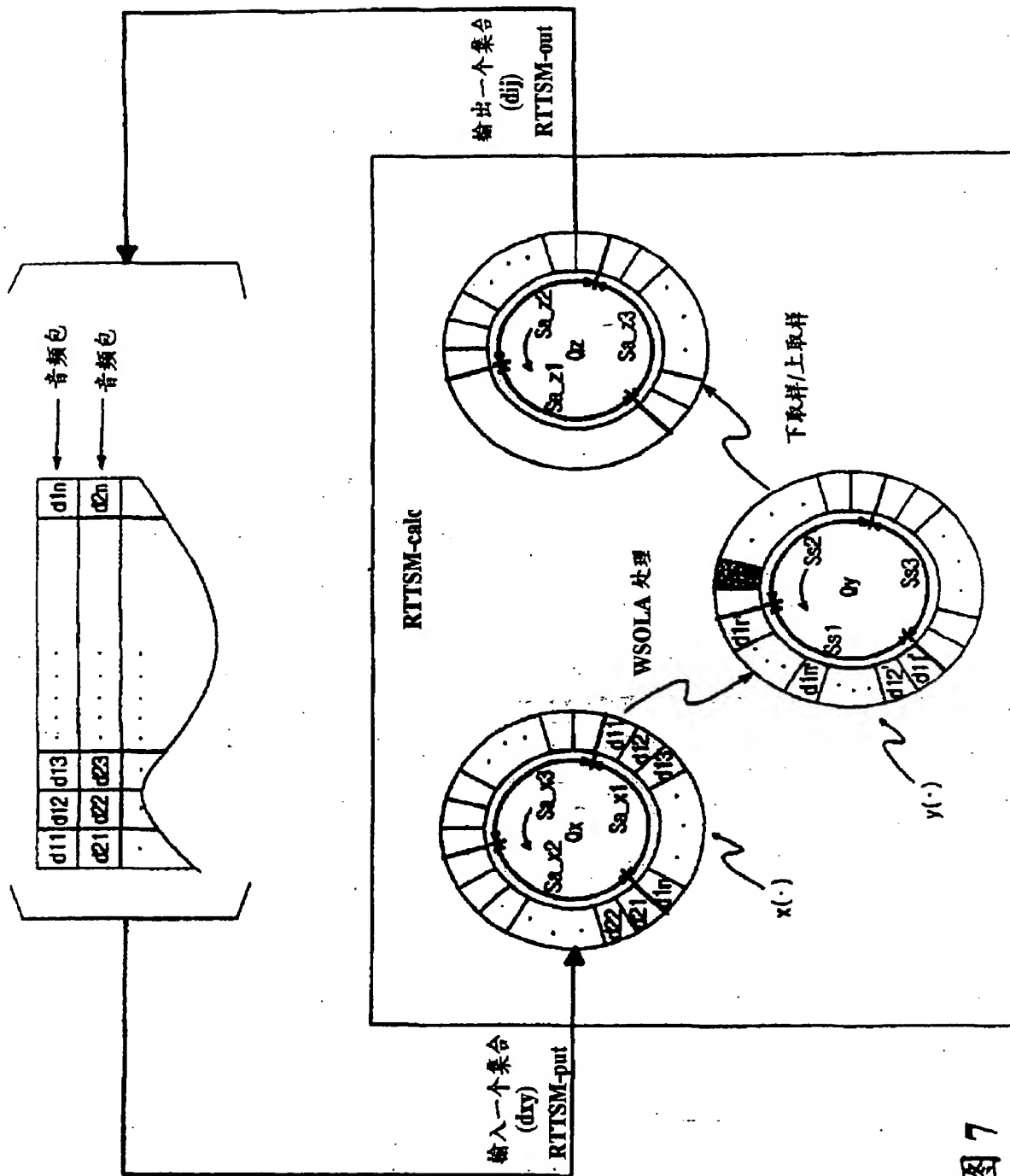


图 7

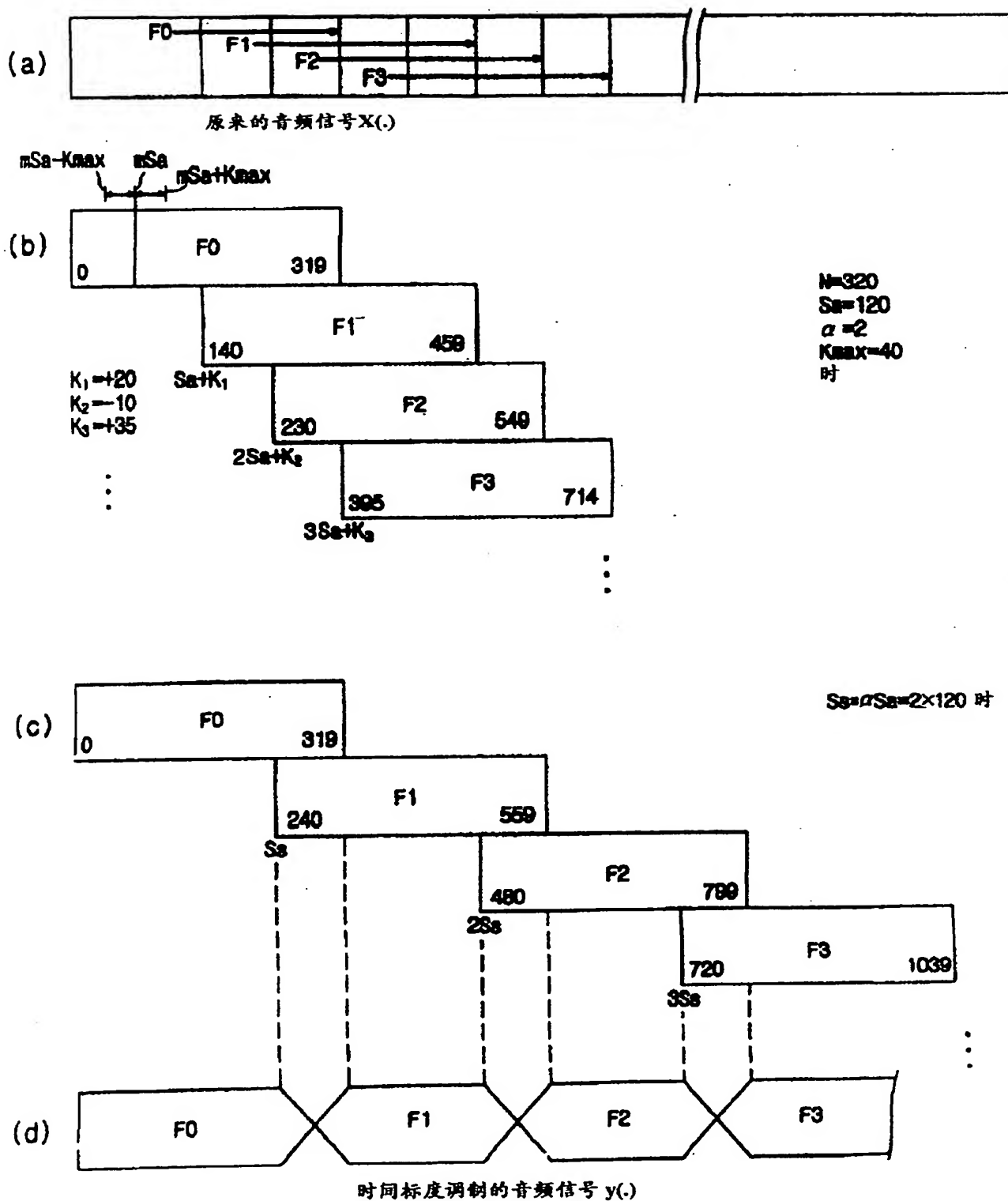


图 8

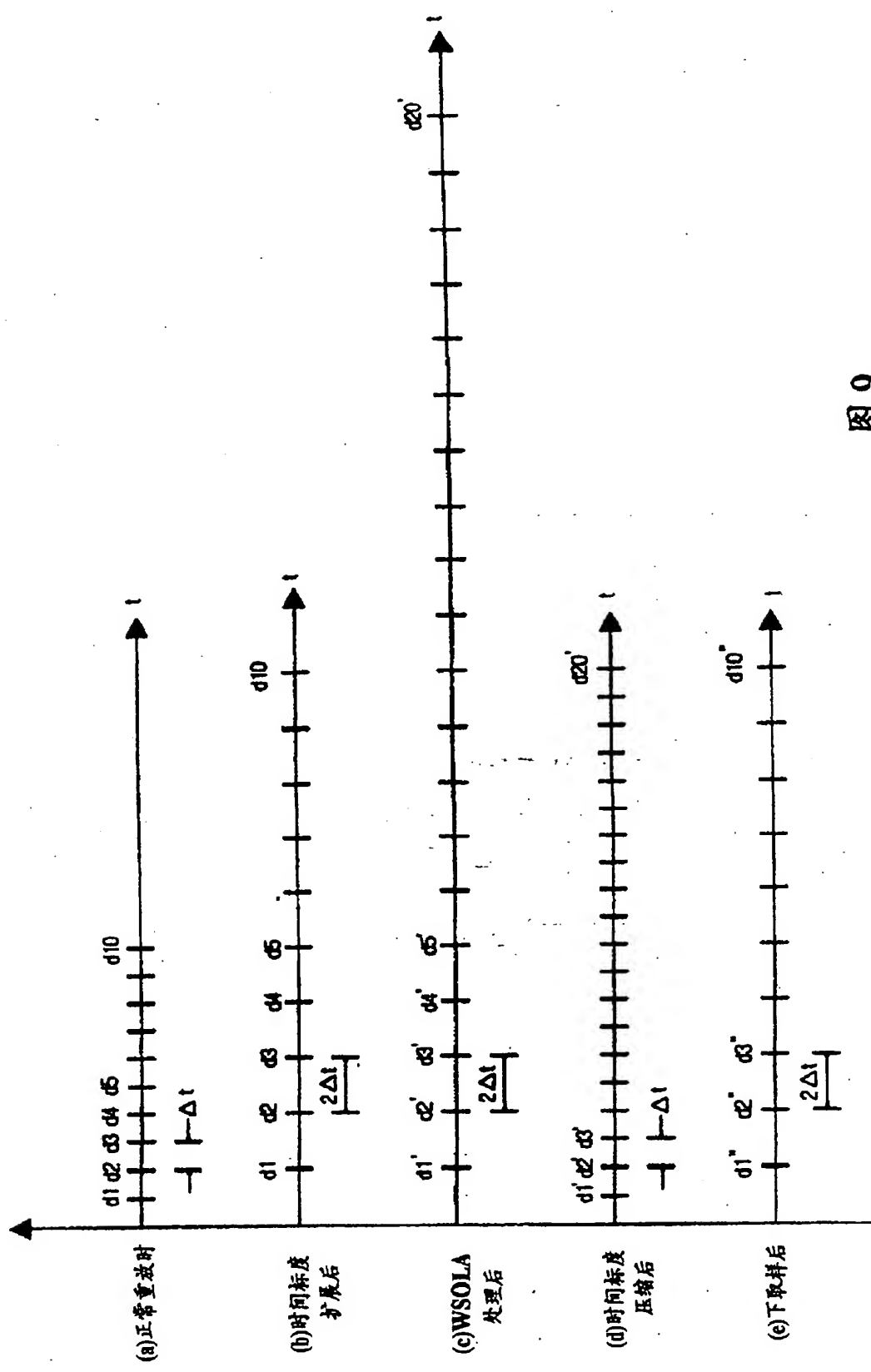


图 9

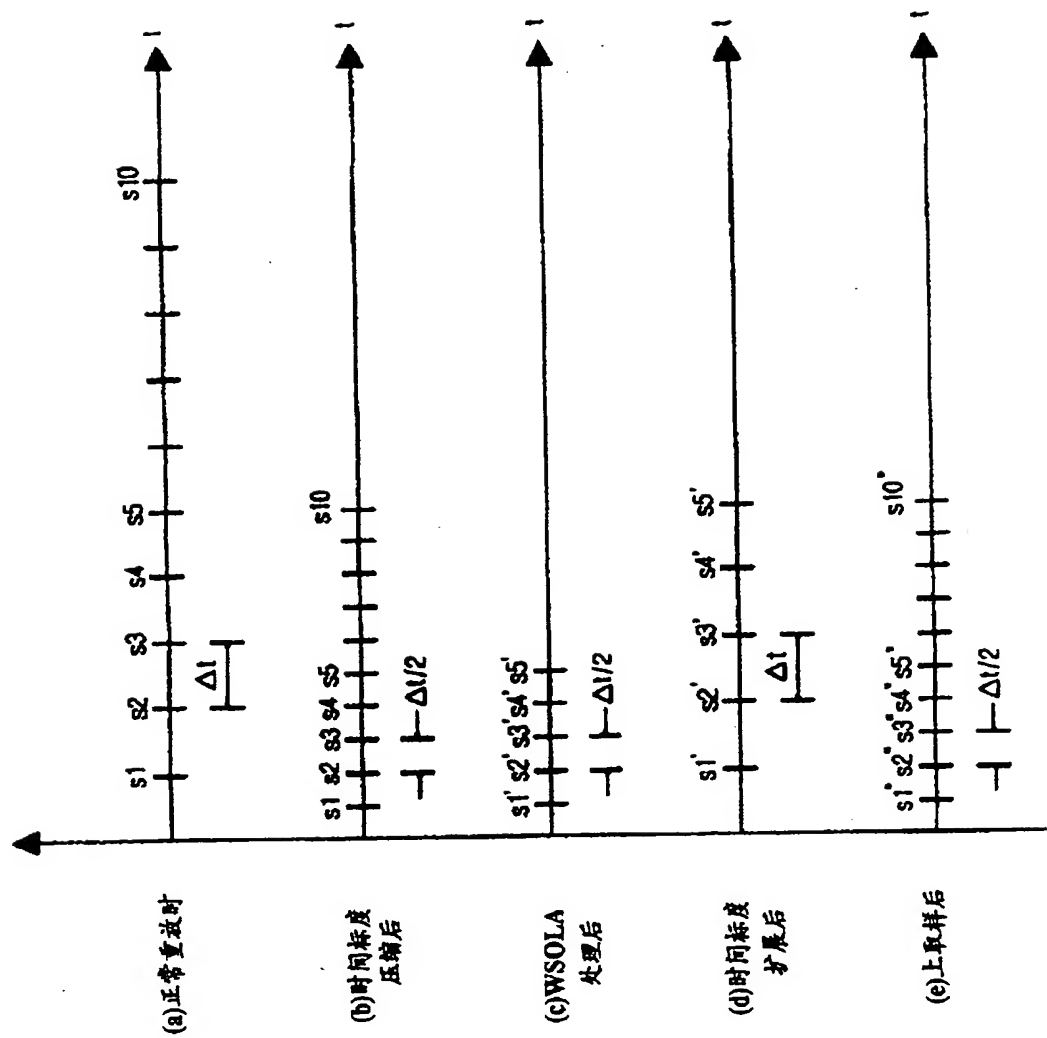
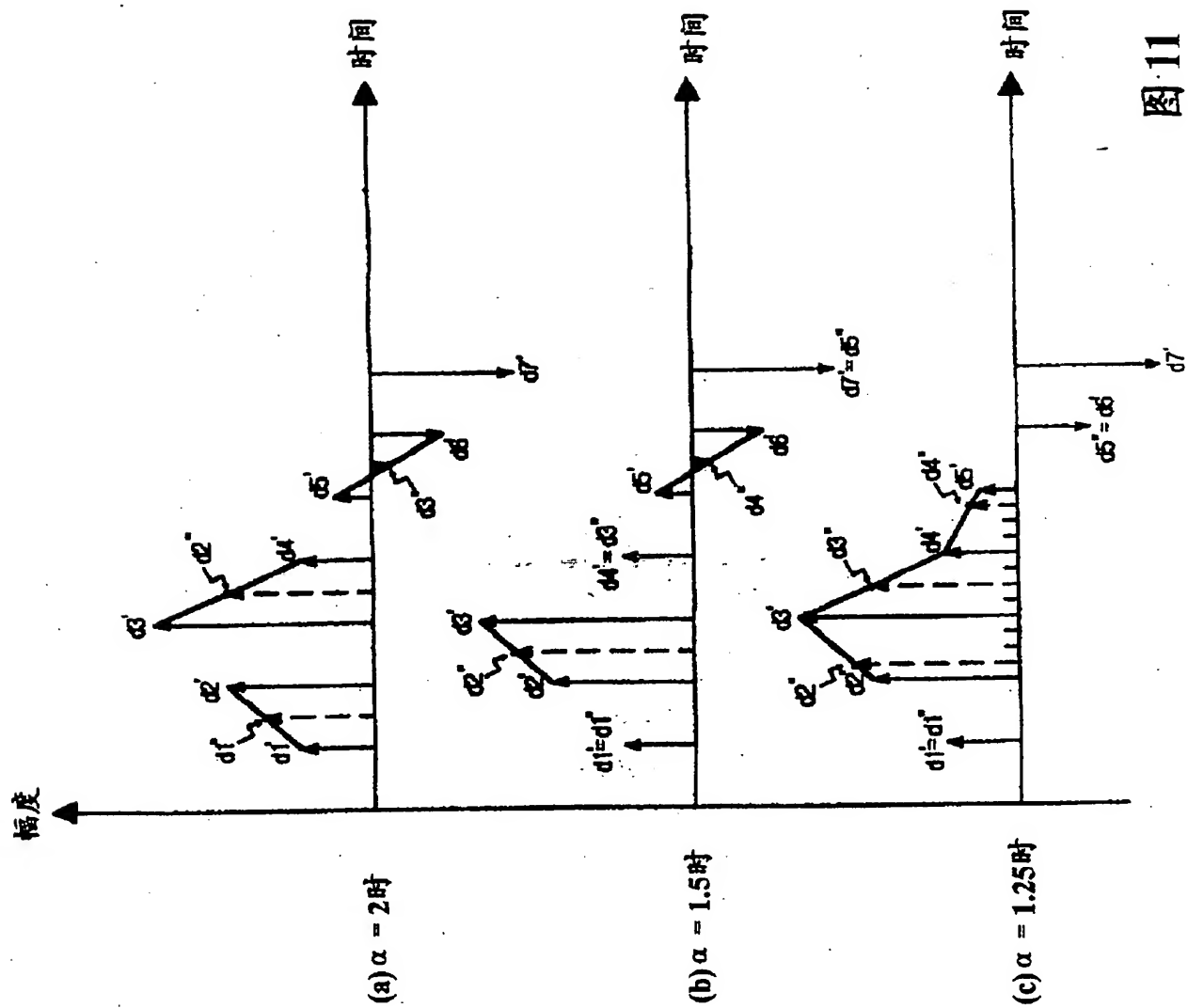


图 10



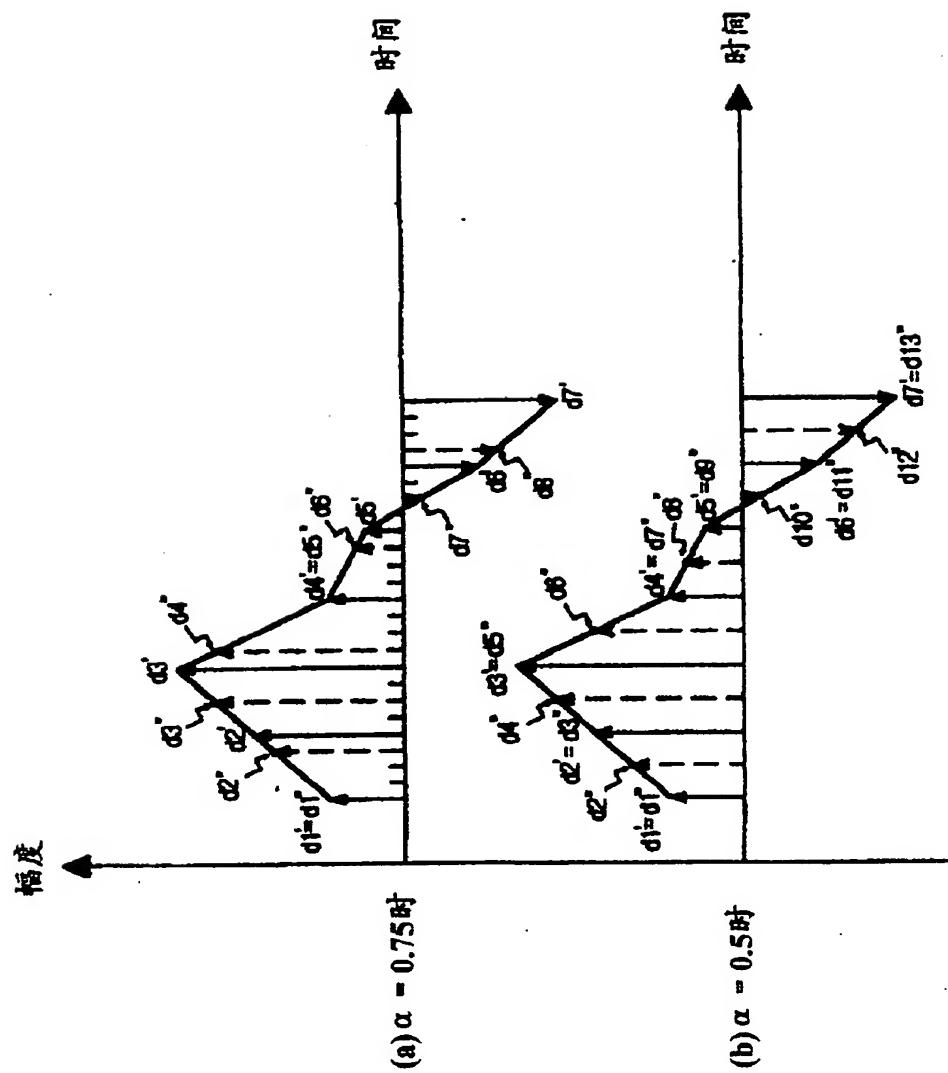


图 12

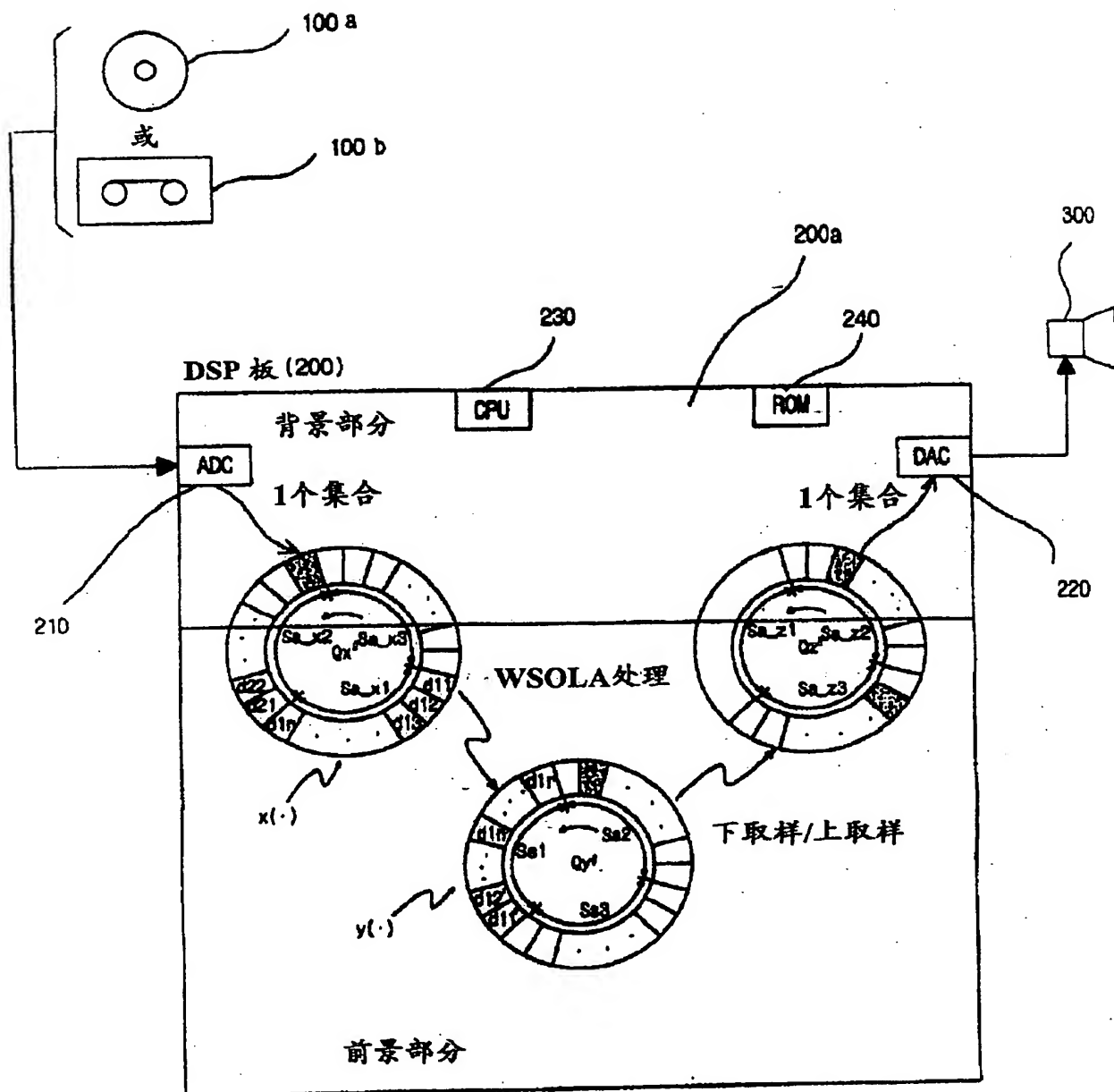


图 13

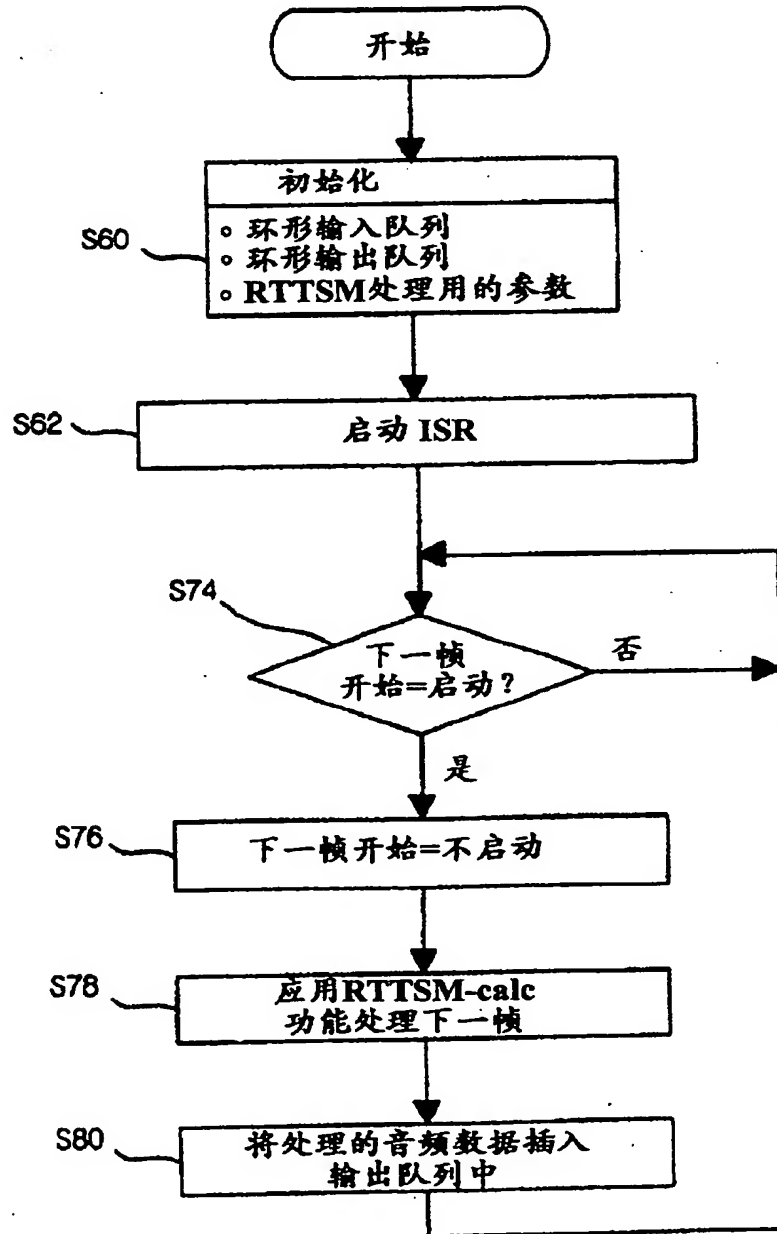


图 14a

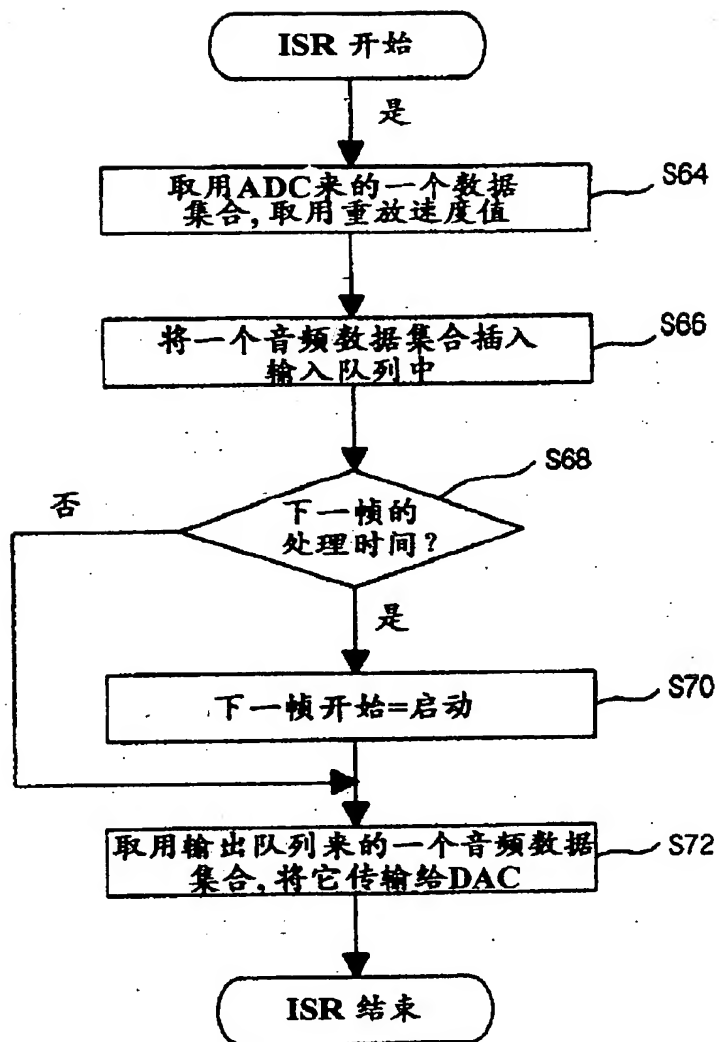


图 14b

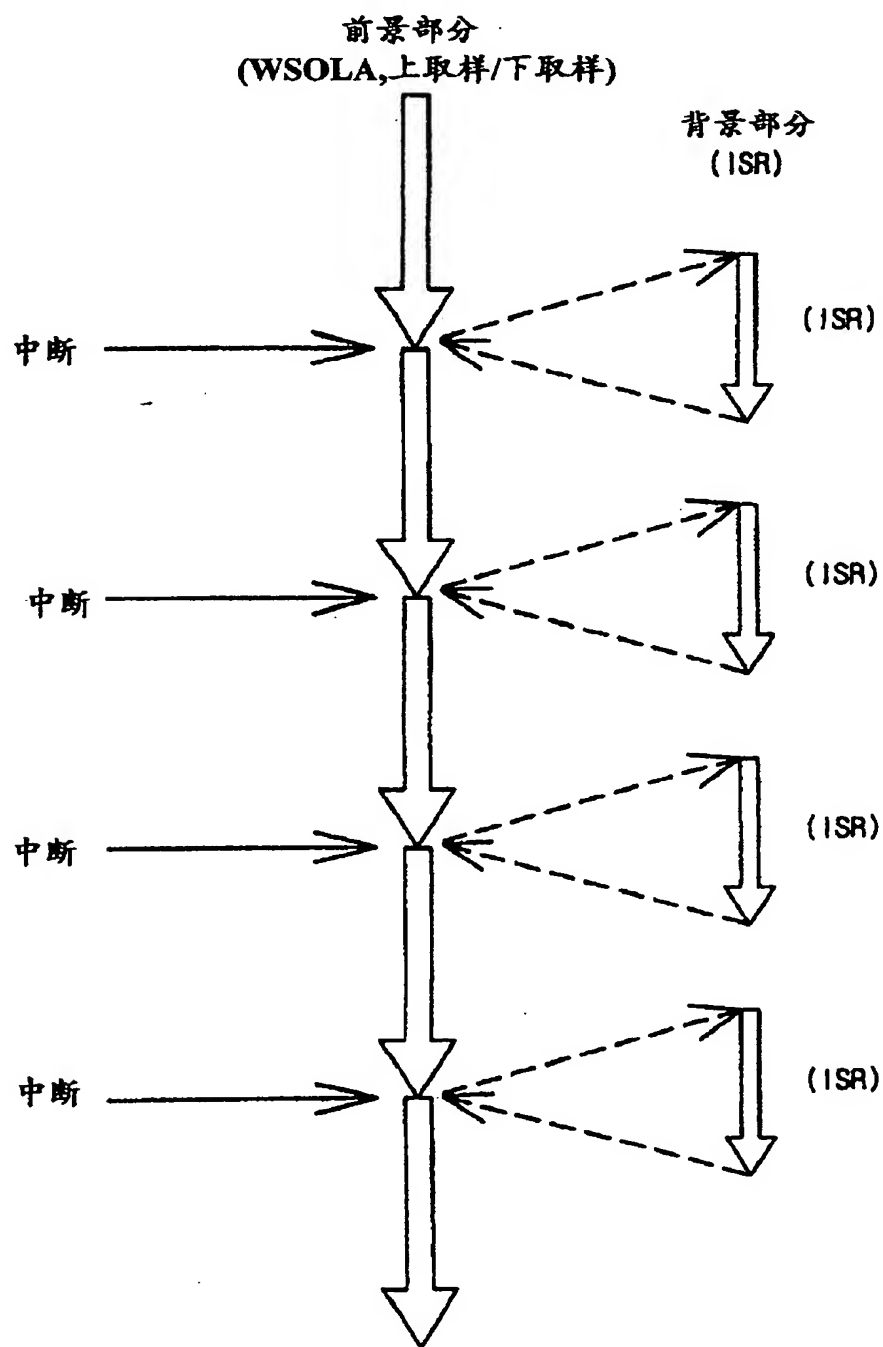


图 14c

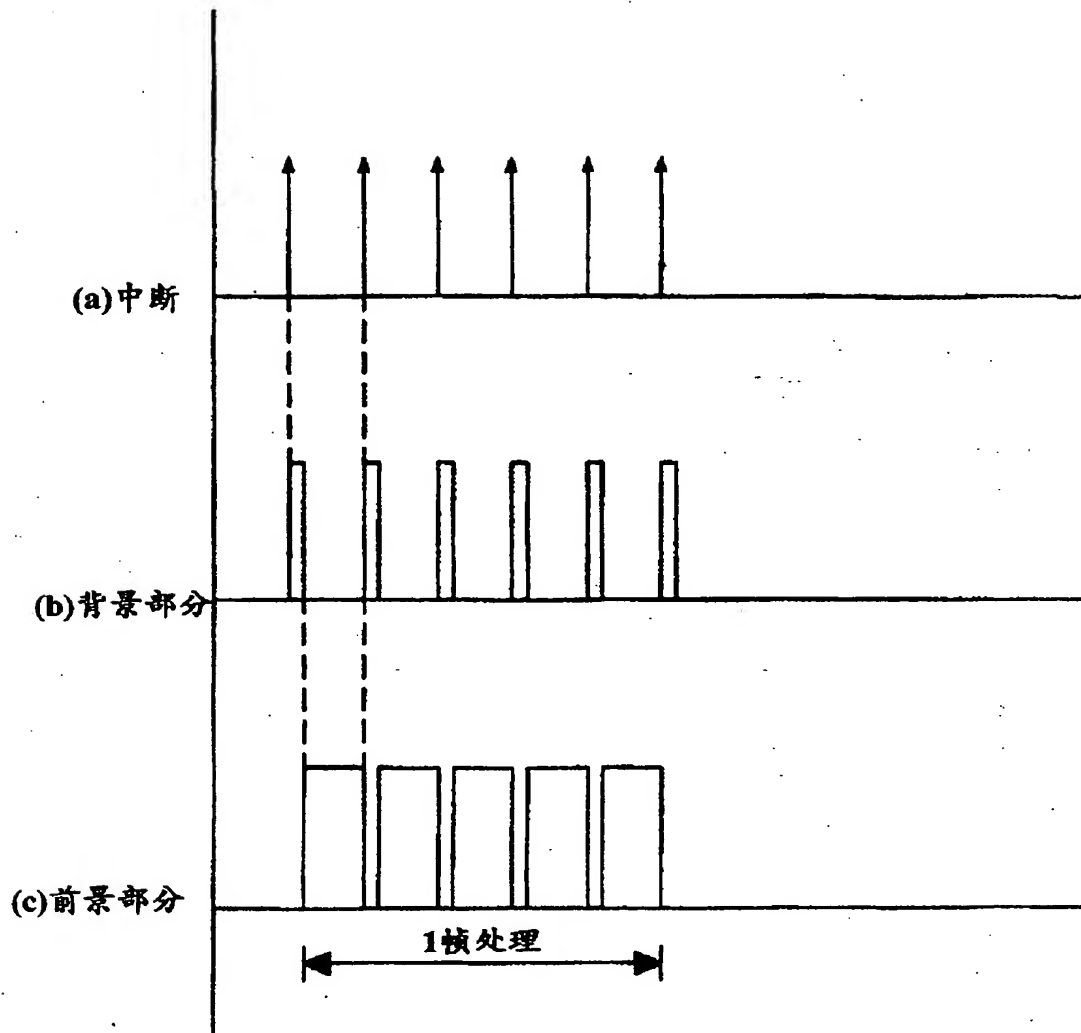


图 15

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☒ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.